GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:25:39 ; Search time 70.371 Seconds
(without alignments)
1758.725 Million cell updates/sec

Title: US-10-622-817-2
Perfect score: 1667
Sequence: 1 MPMYQVKPYHGGGAPLRVEL.....RWMRSCENLAPFNTALKLLK 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1667	100.0	320	7	ADE58613	Ade58613 Human Pro
2	1667	100.0	320	8	ABM80843	Abm80843 Tumour-as
3	1630	97.8	312	2	AAW25776	Aaw25776 JTV1 prot
4	1630	97.8	312	8	ADR86551	Adr86551 1-312 ami
5	1378	82.7	272	7	ADC10204	Adc10204 Human NOV
6	1191	71.4	229	8	ADR86553	Adr86553 84-312 am
7	832	49.9	161	8	ADR86552	Adr86552 1-161 ami
8	247	14.8	51	5	ABP01502	Abp01502 Human ORF
9	214.5	12.9	334	4	ABB62468	Abb62468 Drosophil
10	181	10.9	263	4	ABG23964	Abg23964 Novel hum
11	125.5	7.5	716	3	AAG41411	Aag41411 Arabidops
12	125.5	7.5	719	3	AAG41410	Aag41410 Arabidops
13	125.5	7.5	748	3	AAG41409	Aag41409 Arabidops
14	104	6.2	639	4	ABB62218	Abb62218 Drosophil
15	103	6.2	1512	4	AAU04349	Aau04349 Mammalian
16	103	6.2	1512	4	AAM78732	Aam78732 Human pro
17	103	6.2	1550	4	AAM79716	Aam79716 Human pro
18	102	6.1	201	8	ADS24684	Ads24684 Bacterial
19	101	6.1	559	2	AAR10682	Aar10682 Polyhydro
20	101	6.1	826	4	AAB79740	Aab79740 Corynebac
21	101	6.1	826	4	AAB79989	Aab79989 Corynebac
22	101	6.1	826	4	AAU71908	Aau71908 C. glutam
23	101	6.1	833	4	AAB79739	Aab79739 Corynebac
24	101	6.1	833	4	AAB79988	Aab79988 Corynebac
25	101	6.1	833	4	AAU71907	Aau71907 C. glutam
26	101	6.1	1221	4	AAG91409	Aag91409 C glutami
27	101	6.1	1221	5	ABG80321	Abg80321 C. glutam
28	101	6.1	1221	5	AAE18908	Aae18908 Corynebac
29	101	6.1	1221	7	ADD13589	Add13589 C. glutam
30	100.5	6.0	1200	6	ABU26069	Abu26069 Protein e
31	99.5	6.0	650	2	AAY29614	Aay29614 Mouse PIA
32	99.5	6.0	651	2	AAY29621	Aay29621 Mouse PIA
33	97.5	5.8	217	7	ADE63092	Ade63092 Rat Prote
34	97.5	5.8	217	7	ADD48150	Add48150 Rat Prote
35	97.5	5.8	217	7	ADD48158	Add48158 Rat Prote
36	97.5	5.8	217	7	ADD48154	Add48154 Rat Prote
37	97.5	5.8	217	7	ADE63084	Ade63084 Rat Prote
38	97.5	5.8	217	7	ADE63088	Ade63088 Rat Prote
39	97.5	5.8	218	2	AAR20033	Aar20033 Rat brain
40	97.5	5.8	218	7	ADE56502	Ade56502 Rat Prote
41	97.5	5.8	652	5	ABP41869	Abp41869 Human ova
42	97	5.8	559	3	AAY77178	Aay77178 S. venezu
43	97	5.8	2221	4	ABB61339	Abb61339 Drosophil
44	97	5.8	5215	8	ADL91930	Adl91930 Streptomy
45	96.5	5.8	1049	6	ABM67090	Abm67090 Photorhab

ALIGNMENTS

RESULT 1
ADE58613
ID ADE58613 standard; protein; 320 AA.
XX
AC ADE58613;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein Q13155, SEQ ID NO 4489.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
DR GENBANK; Q13155.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal

CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 320 AA;

Query Match 100.0%; Score 1667; DB 7; Length 320;
Best Local Similarity 100.0%; Pred. No. 5.4e-167;
Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWMRSCENLAPFNTALKLLK 320
              ||||||||||||||||||||
Db        301 RWMRSCENLAPFNTALKLLK 320
```

RESULT 2
ABM80843
ID ABM80843 standard; protein; 320 AA.
XX
AC ABM80843;
XX
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) polypeptide PRO81501, SEQ:2179.
XX
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2004030615-A2.
XX
PD 15-APR-2004.
XX
PF 29-SEP-2003; 2003WO-US028547.
XX
PR 02-OCT-2002; 2002US-0414971P.
XX
PA (GETH) GENENTECH INC.
XX
PI Wu TD, Zhang Z, Zhou Y;
XX
DR WPI; 2004-347921/32.
DR N-PSDB; ACN38579.
XX
PT New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
PS Claim 12; SEQ ID NO 2179; 7273pp; English.
XX
CC The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT polypeptide of the invention
XX
SQ Sequence 320 AA;

Query Match 100.0%; Score 1667; DB 8; Length 320;
Best Local Similarity 100.0%; Pred. No. 5.4e-167;
Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWMRSCENLAPFNTALKLLK 320
              ||||||||||||||||||||
Db        301 RWMRSCENLAPFNTALKLLK 320
```

RESULT 3
AAW25776

ID AAW25776 standard; protein; 312 AA.
XX
AC AAW25776;
XX
DT 19-DEC-1997 (first entry)
XX
DE JTV1 protein.
XX
KW JTV1; hPMS2; probe; detection; chromosome 7; deletion;
KW mismatch repair gene; hereditary non-polyposis colorectal cancer;
KW homologous recombination.
XX
OS Homo sapiens.
XX
PN WO9708312-A1.
XX
PD 06-MAR-1997.
XX
PF 26-AUG-1996; 96WO-US013598.
XX
PR 24-AUG-1995; 95US-00518862.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Vogelstein B, Kinzler KW, Nicolaides NC;
XX
DR WPI; 1997-179269/16.
DR N-PSDB; AAT86182.
XX
PT Novel chromosome 7 gene, JTV1 - used for detecting chromosome 7
PT deletions, and PMS2 promoter activity.
XX
PS Claim 5; Fig 2; 31pp; English.
XX
CC This sequence is JTV1 protein and is encoded by DNA isolated from human
CC chromosome 7. The JTV1 coding sequence is located upstream from hPMS2.
CC JTV1 cDNA can be used as probes to detect chromosome 7 deletions
CC involving JTV1. Due to the overlapping promoter regions, deletions of
CC JTV1 would also affect PMS2 (a mismatch repair gene) expression, leading
CC to hereditary non-polyposis colorectal cancer. JTV1 can also be used to
CC assay activity or competence of the PMS2 promoter region, the presence of
CC JTV1 suggesting that the PMS2 promoter is intact. JTV1 sequences can also
CC be used to guide homologous recombination at the PMS2 locus
XX
SQ Sequence 312 AA;

Query Match 97.8%; Score 1630; DB 2; Length 312;
Best Local Similarity 100.0%; Pred. No. 4.3e-163;
Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWMRSCENLAPF 312
              ||||||||||||
Db        301 RWMRSCENLAPF 312
```

RESULT 4
ADR86551
ID ADR86551 standard; protein; 312 AA.
XX
AC ADR86551;
XX
DT 18-NOV-2004 (first entry)
XX
DE 1-312 amino acid sequence of p38/JTV-1 protein.
XX
KW p38/JTV-1; Cytostatic; cancer; leukemia; anticancer.
XX
OS Homo sapiens.
XX
PN EP1454628-A2.
XX
PD 08-SEP-2004.
XX
PF 09-SEP-2003; 2003EP-00020344.
XX
PR 03-MAR-2003; 2003KR-00013058.
XX
PA (UYSE-) UNIV SEOUL NAT IND FOUND.
XX
PI Kim S, Park B;
XX
DR WPI; 2004-627822/61.
DR N-PSDB; ADR86548.
XX
PT New isolated p38/JTV-1 protein, useful as medicament for treating cancer
PT e.g., stomach, liver, blood, bone, pancreatic, skin, head or neck cancer
PT and cutaneous or intraocular melanoma, as well as for screening new
PT anticancer agents.
XX
PS Claim 5; SEQ ID NO 4; 47pp; English.
XX
CC The present invention relates to an isolated p38/JTV-1 protein for use as
CC medicament. The p38/JTV-1 protein or the pharmaceutical composition is
CC useful as medicament for treating breast cancer, large intestinal cancer,
CC lung cancer, small cell lung cancer, stomach cancer, liver cancer, blood
CC cancer, bone cancer, pancreatic cancer, skin cancer, head or neck cancer,
CC cutaneous or intraocular melanoma, uterine sarcoma, ovarian cancer,
CC rectal cancer, anal cancer, colon cancer, fallopian tube carcinoma,
CC endometrial carcinoma, cervical cancer, vulval cancer, vaginal carcinoma,
CC Hodgkin's disease, esophageal cancer, small intestine cancer, endocrine
CC cancer, thyroid cancer, parathyroid cancer, adrenal cancer, soft tissue
CC tumour, urethral cancer, penile cancer, prostate cancer, chronic or acute
CC leukemia, lymphocytic lymphoma, bladder cancer, kidney cancer, ureter
CC cancer, renal cell carcinoma, renal pelvic carcinoma, CNS tumour, primary
CC CNS lymphoma, bone marrow tumour, brain stem nerve gliomas, pituitary
CC adenoma, or their combination. The protein is useful as a target for
CC screening new anticancer agents. The present sequence represents the 1-
CC 312 amino acid sequence of p38/JTV-1 protein.
XX
SQ Sequence 312 AA;

Query Match 97.8%; Score 1630; DB 8; Length 312;
Best Local Similarity 100.0%; Pred. No. 4.3e-163;
Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
```

```
Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWMRSCENLAPF 312
              ||||||||||||
Db        301 RWMRSCENLAPF 312
```

RESULT 5
ADC10204
ID ADC10204 standard; protein; 272 AA.
XX
AC ADC10204;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human NOVX polypeptide SEQ ID NO: 226.
XX
KW cytostatic; antidiabetic; anorectic; cerebroprotective; neuroprotective;
KW antiinflammatory; gene therapy; antisense therapy; thyromimetic; NOVX;
KW pathology; cancer; diabetes; obesity; endocrine disorder; CNS disorder;
KW inflammatory disorder; chromosome mapping; tissue typing;
KW predictive medicine.
XX
OS Homo sapiens.
XX
PN WO2003000842-A2.
XX
PD 03-JAN-2003.
XX
PF 04-JUN-2002; 2002WO-US017443.
XX
PR 04-JUN-2001; 2001US-0295607P.
PR 04-JUN-2001; 2001US-0295661P.
PR 06-JUN-2001; 2001US-0296404P.
PR 06-JUN-2001; 2001US-0296418P.
PR 07-JUN-2001; 2001US-0296575P.
PR 11-JUN-2001; 2001US-0297414P.
PR 12-JUN-2001; 2001US-0295573P.
PR 12-JUN-2001; 2001US-0297567P.
PR 14-JUN-2001; 2001US-0298285P.
PR 15-JUN-2001; 2001US-0298528P.
PR 18-JUN-2001; 2001US-0299133P.
PR 19-JUN-2001; 2001US-0299230P.
PR 21-JUN-2001; 2001US-0299949P.
PR 22-JUN-2001; 2001US-0300177P.
PR 26-JUN-2001; 2001US-0300883P.
PR 28-JUN-2001; 2001US-0301530P.
PR 28-JUN-2001; 2001US-0301550P.
PR 03-JUL-2001; 2001US-0302951P.
PR 31-JUL-2001; 2001US-0308890P.
PR 14-SEP-2001; 2001US-0322297P.
PR 25-SEP-2001; 2001US-0324669P.
PR 03-DEC-2001; 2001US-0337477P.
PR 14-DEC-2001; 2001US-0341562P.
PR 21-FEB-2002; 2002US-0358656P.
PR 21-FEB-2002; 2002US-0359122P.
PR 22-FEB-2002; 2002US-0358978P.
PR 22-FEB-2002; 2002US-0359034P.
PR 22-FEB-2002; 2002US-0359035P.
PR 22-FEB-2002; 2002US-0359121P.
PR 27-FEB-2002; 2002US-0359964P.
PR 01-MAR-2002; 2002US-0360858P.
PR 12-MAR-2002; 2002US-0363430P.
PR 12-MAR-2002; 2002US-0363676P.
PR 10-APR-2002; 2002US-0371346P.
PR 10-MAY-2002; 2002US-0379444P.
PR 04-JUN-2002; 2002US-00379444.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Agee ML, Anderson DW, Berghs C, Casman SJ, Catterton E;
PI Dipippo VA, Edinger SR, Eisen A, Ellerman K, Gangolli EA;
PI Gerlach VL, Gorman L, Guo X, Herrmann JL, Hjalt T, Ji W, Kekuda R;
PI Khramtsov NV, Li L, Liu X, Malyankar UM, Miller CE, Millet I;
PI Ort T, Padigaru M, Patturajan M, Pena CEA, Rastelli L, Rieger DK;
PI Rothenberg ME, Shenoy SG, Shimkets RA, Smithson G, Spaderna SK;
PI Spytek KA, Stone DJ, Vernet CAM, Zhong H, Zhong M, Alsobrook JP;
PI Burgess CE, Lepley DM;
XX
DR WPI; 2003-210149/20.
DR N-PSDB; ADC10203.
XX
PT New isolated NOVX polypeptides and nucleic acid molecules useful for
PT treating, preventing and diagnosing pathological conditions with NOVX-
PT associated disorders, such as cancer, obesity, diabetes and inflammatory
PT or CNS diseases.
XX
PS Claim 1; SEQ ID NO 226; 772pp; English.
XX
CC The invention relates to novel isolated polypeptides, mature form of the
CC polypeptide, a sequence that is 95% identical to the polypeptide or the
CC polypeptide comprising one or more conservative substitutions. The NOVX
CC polypeptide is useful for treating or preventing a pathology associated
CC with the polypeptide e.g. disorders associated with aberrant expression
CC or activity of the polypeptide, such as cancer, diabetes, obesity, and
CC endocrine, CNS and inflammatory disorders. They can also be used in
CC various detection and screening assays, chromosome mapping, tissue typing
CC and predictive medicine. This sequence corresponds to one of the
CC polypeptides of the invention.
XX
SQ Sequence 272 AA;

Query Match 82.7%; Score 1378; DB 7; Length 272;
Best Local Similarity 100.0%; Pred. No. 1.6e-136;
Matches 268; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         45 QEESNLSLQALESRQDDILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTN 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 QEESNLSLQALESRQDDILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTN 64

Qy        105 ALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPEN 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 ALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPEN 124

Qy        165 LLKCFGEQNKKQPRQDYQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQK 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 LLKCFGEQNKKQPRQDYQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQK 184

Qy        225 HNAVNATLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSV 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 HNAVNATLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSV 244

Qy        285 LQQIGGCSVTVPANVQRWMRSCENLAPF 312
              ||||||||||||||||||||||||||||
Db        245 LQQIGGCSVTVPANVQRWMRSCENLAPF 272
```

RESULT 6
ADR86553
ID ADR86553 standard; protein; 229 AA.
XX
AC ADR86553;
XX
DT 18-NOV-2004 (first entry)
XX
DE 84-312 amino acid sequence of p38/JTV-1 protein.
XX
KW p38/JTV-1; Cytostatic; cancer; leukemia; anticancer.
XX

OS Homo sapiens.
XX
PN EP1454628-A2.
XX
PD 08-SEP-2004.
XX
PF 09-SEP-2003; 2003EP-00020344.
XX
PR 03-MAR-2003; 2003KR-00013058.
XX
PA (UYSE-) UNIV SEOUL NAT IND FOUND.
XX
PI Kim S, Park B;
XX
DR WPI; 2004-627822/61.
DR N-PSDB; ADR86550.
XX
PT New isolated p38/JTV-1 protein, useful as medicament for treating cancer
PT e.g., stomach, liver, blood, bone, pancreatic, skin, head or neck cancer
PT and cutaneous or intraocular melanoma, as well as for screening new
PT anticancer agents.
XX
PS Claim 5; SEQ ID NO 6; 47pp; English.
XX
CC The present invention relates to an isolated p38/JTV-1 protein for use as
CC medicament. The p38/JTV-1 protein or the pharmaceutical composition is
CC useful as medicament for treating breast cancer, large intestinal cancer,
CC lung cancer, small cell lung cancer, stomach cancer, liver cancer, blood
CC cancer, bone cancer, pancreatic cancer, skin cancer, head or neck cancer,
CC cutaneous or intraocular melanoma, uterine sarcoma, ovarian cancer,
CC rectal cancer, anal cancer, colon cancer, fallopian tube carcinoma,
CC endometrial carcinoma, cervical cancer, vulval cancer, vaginal carcinoma,
CC Hodgkin's disease, esophageal cancer, small intestine cancer, endocrine
CC cancer, thyroid cancer, parathyroid cancer, adrenal cancer, soft tissue
CC tumour, urethral cancer, penile cancer, prostate cancer, chronic or acute
CC leukemia, lymphocytic lymphoma, bladder cancer, kidney cancer, ureter
CC cancer, renal cell carcinoma, renal pelvic carcinoma, CNS tumour, primary
CC CNS lymphoma, bone marrow tumour, brain stem nerve gliomas, pituitary
CC adenoma, or their combination. The protein is useful as a target for
CC screening new anticancer agents. The present sequence represents the 84-
CC 312 amino acid sequence of p38/JTV-1 protein.
XX
SQ Sequence 229 AA;

Query Match 71.4%; Score 1191; DB 8; Length 229;
Best Local Similarity 100.0%; Pred. No. 7.2e-117;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         84 DADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLC 143
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 DADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLC 60

Qy        144 EHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQDYQLGFTLIWKNVPKTQMKFSIQT 203
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQDYQLGFTLIWKNVPKTQMKFSIQT 120

Qy        204 MCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSAL 263
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 MCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSAL 180

Qy        264 GKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQRWMRSCENLAPF 312
               |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQRWMRSCENLAPF 229
```

RESULT 7
ADR86552
ID ADR86552 standard; protein; 161 AA.
XX
AC ADR86552;
XX
DT 18-NOV-2004 (first entry)
XX
DE 1-161 amino acid sequence of p38/JTV-1 protein.
XX
KW p38/JTV-1; Cytostatic; cancer; leukemia; anticancer.
XX
OS Homo sapiens.
XX
PN EP1454628-A2.
XX
PD 08-SEP-2004.
XX
PF 09-SEP-2003; 2003EP-00020344.
XX
PR 03-MAR-2003; 2003KR-00013058.
XX
PA (UYSE-) UNIV SEOUL NAT IND FOUND.
XX
PI Kim S, Park B;
XX
DR WPI; 2004-627822/61.
DR N-PSDB; ADR86549.
XX
PT New isolated p38/JTV-1 protein, useful as medicament for treating cancer
PT e.g., stomach, liver, blood, bone, pancreatic, skin, head or neck cancer
PT and cutaneous or intraocular melanoma, as well as for screening new
PT anticancer agents.
XX
PS Claim 5; SEQ ID NO 5; 47pp; English.
XX
CC The present invention relates to an isolated p38/JTV-1 protein for use as
CC medicament. The p38/JTV-1 protein or the pharmaceutical composition is
CC useful as medicament for treating breast cancer, large intestinal cancer,
CC lung cancer, small cell lung cancer, stomach cancer, liver cancer, blood
CC cancer, bone cancer, pancreatic cancer, skin cancer, head or neck cancer,
CC cutaneous or intraocular melanoma, uterine sarcoma, ovarian cancer,
CC rectal cancer, anal cancer, colon cancer, fallopian tube carcinoma,
CC endometrial carcinoma, cervical cancer, vulval cancer, vaginal carcinoma,
CC Hodgkin's disease, esophageal cancer, small intestine cancer, endocrine
CC cancer, thyroid cancer, parathyroid cancer, adrenal cancer, soft tissue
CC tumour, urethral cancer, penile cancer, prostate cancer, chronic or acute
CC leukemia, lymphocytic lymphoma, bladder cancer, kidney cancer, ureter
CC cancer, renal cell carcinoma, renal pelvic carcinoma, CNS tumour, primary
CC CNS lymphoma, bone marrow tumour, brain stem nerve gliomas, pituitary
CC adenoma, or their combination. The protein is useful as a target for
CC screening new anticancer agents. The present sequence represents the 1-
CC 161 amino acid sequence of p38/JTV-1 protein.
XX
SQ Sequence 161 AA;

Query Match 49.9%; Score 832; DB 8; Length 161;
Best Local Similarity 100.0%; Pred. No. 3.9e-79;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSV 161
              |||||||||||||||||||||||||||||||||||||||||
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSV 161
```

RESULT 8
ABP01502
ID ABP01502 standard; protein; 51 AA.
XX
AC ABP01502;
XX
DT 24-JUN-2002 (first entry)
XX

```
DE   Human ORFX protein sequence SEQ ID NO:2986.
XX
KW   Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW   hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW   degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW   cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW   hypertension; hypothyroidism; cholesterol ester storage disease;
KW   immune deficiency; immune disorder; infectious disease;
KW   autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW   myasthenia gravis.
XX
OS   Homo sapiens.
XX
PN   WO200192523-A2.
XX
PD   06-DEC-2001.
XX
PF   29-MAY-2001; 2001WO-US010836.
XX
PR   30-MAY-2000; 2000US-0206132P.
PR   29-AUG-2000; 2000US-0228716P.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach MD;
XX
DR   WPI; 2002-106308/14.
DR   N-PSDB; ABN17254.
XX
PT   Novel human polypeptides and polynucleotides useful for diagnosing,
PT   preventing and treating cardiovascular disease, neurodegenerative,
PT   hyperproliferative disorders and autoimmune disorders.
XX
PS   Disclosure; SEQ ID NO 2986; 1037pp; English.
XX
CC   The present invention describes substantially purified human proteins
CC   (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC   in the specification). ABN15762 to ABN27252 encode the human ORFX
CC   proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC   treating or preventing a pathology associated with an ORFX-associated
CC   disorder in humans, and in the manufacture of a medicament for treating a
CC   syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC   sequences can be used in gene therapy. ORFX sequences can be used in the
CC   treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC   psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC   osteoarthritis, neurodegenerative disorders, disorders related to organ
CC   transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC   lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC   storage disease, various immune deficiencies and disorders, infectious
CC   diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC   arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC   disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC   useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC   bone degenerative disorders, or periodontal disease, and for gut
CC   protection or regeneration and treatment of lung or liver fibrosis,
CC   reperfusion injury in various tissues and conditions resulting from
CC   systemic cytokine damage. N.B. The sequence data for this patent did not
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 51 AA;

  Query Match             14.8%;  Score 247;  DB 5;  Length 51;
  Best Local Similarity   96.1%;  Pred. No. 4.5e-18;
  Matches   49;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         74 DGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALKDIVI 124
              ||||||| ||||||||||:|||||||||||||||||||||||||||||||
Db          1 DGLSKMIHTPDADLDVTNILQADEPTTLTTNALDLNSVLGKDYGALKDIVI 51

RESULT 9
ABB62468
ID   ABB62468 standard; protein; 334 AA.
XX
AC   ABB62468;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 14196.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US009231.
XX
PR   23-MAR-2000; 2000US-0191637P.
PR   11-JUL-2000; 2000US-00614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL06571.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signaling and cell-cell
PT   interactions.
XX
PS   Disclosure; SEQ ID NO 14196; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC   ABB72072). The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 334 AA;

  Query Match             12.9%;  Score 214.5;  DB 4;  Length 334;
  Best Local Similarity   25.6%;  Pred. No. 2.8e-13;
  Matches   85;  Conservative   50;  Mismatches  120;  Indels   77;  Gaps   10;

Qy         18 VELPTCMYRLPNV----HGRSYGPAPGAGHVQEESNLSLQA------------------ 54
              ::|||||| | ||       : | :  |      |:  |:|
Db         13 IKLPTCMYPLKNVSLAADSLASGSSTSASTSASTSSCKLEANRIDRTGRNAATCALDLDS 72

Qy         55 --------------LESRQDDILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPT 99
                           : :||: :||:| |||| :  :          | | |    : : |
Db         73 LGRQIQRLLKDDTASVAARQEKVLKQLEELKAQLGQIR-------AGLGVCG--KTFQHT 123

Qy        100 TLTTNALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVK 159
              |  |         |      |:|:||| :|   | :|| |          : :     |||::
Db        124 TAFQNG-------GLKEVPLQDVVINGHPNFIPYALLALKNAWRNLYTIDVKTFTHSTMA 176

Qy        160 SV-------PENLLKCFGEQNKKQPRQDYQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGN 212
              :          || |      |  :    : ||||||    |:|  |      || ||| |
Db        177 DIGPAAREFEANLAKV--PVNPALPK----ISVTLIWKNCEHTEMISSPTMYVPIYGEVN 230

Qy        213 IARFLFSLFGQKHNAVNATL---IDSWVDIAIFQLKEGSSKEKAAVFRSMNSALGKSPWL 269
              | |:|  :   ::    : |   ||  :||     |:  : | : |: | ::  | |  :
Db        231 IIRYLGRVGPAEYRYEGSPLCNEIDLVLDICYQLLRCNTHKTQVAMVRLLDKRLQKQQYF 290

Qy        270 AGNELTVADVVLWSVLQQIGGCSVTVPANVQR 301
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:36:20 ; Search time 17.81 Seconds
(without alignments)
1341.256 Million cell updates/sec

Title: US-10-622-817-2
Perfect score: 1667
Sequence: 1 MPMYQVKPYHGGGAPLRVEL.....RWMRSCENLAPFNTALKLLK 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1667	100.0	341	4	US-09-949-016-11312	Sequence 11312, A
2	1630	97.8	312	2	US-08-518-862C-2	Sequence 2, Appli
3	103	6.2	1512	3	US-09-443-184-48	Sequence 48, Appl
4	97	5.8	559	2	US-08-756-317-10	Sequence 10, Appl
5	97	5.8	559	4	US-09-091-609-4	Sequence 4, Appli
6	97	5.8	5215	3	US-09-105-537-2	Sequence 2, Appli
7	95	5.7	559	4	US-09-821-016-1	Sequence 1, Appli
8	95	5.7	559	4	US-10-259-632-1	Sequence 1, Appli
9	95	5.7	559	4	US-10-266-787-1	Sequence 1, Appli
10	92	5.5	2954	4	US-09-150-867-1	Sequence 1, Appli
11	91	5.5	474	4	US-09-248-796A-20321	Sequence 20321, A
12	90.5	5.4	359	3	US-09-540-824-2	Sequence 2, Appli
13	90.5	5.4	597	1	US-08-399-696-102	Sequence 102, App
14	88.5	5.3	1007	4	US-09-538-092-736	Sequence 736, App
15	88	5.3	427	4	US-09-949-016-11178	Sequence 11178, A
16	88	5.3	724	4	US-09-900-920-62	Sequence 62, Appl
17	88	5.3	883	4	US-09-976-239-2	Sequence 2, Appli
18	88	5.3	883	4	US-10-289-779B-2	Sequence 2, Appli
19	88	5.3	914	4	US-09-976-239-4	Sequence 4, Appli
20	88	5.3	914	4	US-10-289-779B-4	Sequence 4, Appli
21	86.5	5.2	761	3	US-09-012-710-13	Sequence 13, Appl
22	86.5	5.2	761	3	US-09-556-273-13	Sequence 13, Appl
23	86	5.2	211	4	US-09-328-352-6834	Sequence 6834, Ap
24	86	5.2	351	4	US-08-178-257-6	Sequence 6, Appli
25	86	5.2	471	3	US-09-134-001C-4904	Sequence 4904, Ap
26	86	5.2	1365	3	US-09-376-330-18	Sequence 18, Appl
27	86	5.2	1365	6	5194600-4	Patent No. 5194600
28	86	5.2	1365	6	5194600-4	Patent No. 5194600
29	85.5	5.1	1440	3	US-09-357-251-37	Sequence 37, Appl
30	85	5.1	559	4	US-10-253-509-1	Sequence 1, Appli
31	85	5.1	1135	2	US-08-469-537A-97	Sequence 97, Appl
32	84.5	5.1	579	4	US-09-107-433-4488	Sequence 4488, Ap
33	84.5	5.1	829	1	US-07-670-611-2	Sequence 2, Appli
34	84.5	5.1	829	1	US-08-220-674-2	Sequence 2, Appli
35	84.5	5.1	829	1	US-08-445-186-2	Sequence 2, Appli
36	84.5	5.1	829	1	US-08-446-549-2	Sequence 2, Appli
37	84.5	5.1	829	2	US-08-446-550-2	Sequence 2, Appli
38	83.5	5.0	778	4	US-09-583-110-3930	Sequence 3930, Ap
39	83.5	5.0	786	4	US-09-107-433-3893	Sequence 3893, Ap
40	83	5.0	559	2	US-09-052-339-1	Sequence 1, Appli
41	83	5.0	559	3	US-09-385-742B-1	Sequence 1, Appli
42	83	5.0	559	4	US-09-989-786-1	Sequence 1, Appli
43	83	5.0	767	3	US-08-836-567-8	Sequence 8, Appli
44	83	5.0	767	4	US-09-606-304-8	Sequence 8, Appli
45	82.5	4.9	222	4	US-09-248-796A-18146	Sequence 18146, A

ALIGNMENTS

RESULT 1
US-09-949-016-11312
; Sequence 11312, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11312
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11312

Query Match 100.0%; Score 1667; DB 4; Length 341;
Best Local Similarity 100.0%; Pred. No. 2.4e-181;
Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 81

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 141

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 201

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 261

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 321

Qy        301 RWMRSCENLAPFNTALKLLK 320
```

```
                    ||||||||||||||||||||
Db        322 RWMRSCENLAPFNTALKLLK 341


RESULT 2
US-08-518-862C-2
; Sequence 2, Application US/08518862C
; Patent No. 5843757
;  GENERAL INFORMATION:
;    APPLICANT:  Vogelstein, Bert
;    APPLICANT:  Kinzler, Kenneth W.
;    APPLICANT:  Nicolaides, Nicholas C.
;    TITLE OF INVENTION:  Human JTV1 Gene Overlaps PMS2 Gene
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Witcoff, Ltd.
;      STREET:  1001 G Street, N.W.
;      CITY:  Washington, D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/518,862C
;      FILING DATE:  24-AUG-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kagan, Sarah A.
;      REGISTRATION NUMBER:  32,141
;      REFERENCE/DOCKET NUMBER:  01107.49697
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-508-9100
;      TELEFAX:  202-508-9299
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  312 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-518-862C-2

  Query Match             97.8%;  Score 1630;  DB 2;  Length 312;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-177;
  Matches  312;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWMRSCENLAPF 312
              ||||||||||||
Db        301 RWMRSCENLAPF 312


RESULT 3
US-09-443-184-48
; Sequence 48, Application US/09443184A
; Patent No. 6372431
; GENERAL INFORMATION:
;  APPLICANT: Cunningham, Mary Jane
;  APPLICANT: Zweiger, Gary
;  APPLICANT: Kaser, Matthew R.
;  APPLICANT: Panzer, Scott
;  APPLICANT: Seilhammer, Jeffrey J.
;  APPLICANT: Yue, Henry
;  APPLICANT: Baughn, Mariah
;  APPLICANT: Azimzai, Yalda
;  APPLICANT: Lal, Preeti
;  TITLE OF INVENTION: MAMMALIAN TOXICOLOGICAL RESPONSE MARKERS
;  FILE REFERENCE: PC-0007 US
;  CURRENT APPLICATION NUMBER: US/09/443,184A
;  CURRENT FILING DATE: 1999-11-19
;  NUMBER OF SEQ ID NOS: 138
;  SOFTWARE: PERL Program
; SEQ ID NO 48
;   LENGTH: 1512
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6372431 2302721CD1
US-09-443-184-48

  Query Match             6.2%;  Score 103;  DB 3;  Length 1512;
  Best Local Similarity   22.1%;  Pred. No. 0.12;
  Matches   45;  Conservative   30;  Mismatches   69;  Indels   60;  Gaps    9;

Qy        124 INANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQDYQL 183
              :  |   |||  |    |  || :   ::    ||:   ||:|
Db          6 LTVNSGDPPLGAL----LAVEHVKDDVSI----SVEEGKENILH---------------- 41

Qy        184 GFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFL------FSLFGQKHNAVNATLIDSWV 237
              : :||  | :              :| |:|         |:|   | :  | || |:
Db         42 ----VSENVIFTDV------------NSILRYLARVATTAGLYGS--NLMEHTEIDHWL 82

Qy        238 DIAIFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCS----- 292
              : :  :|   |     :    :|  |    :| || |::||: :|: |:  |  :
Db         83 EFSATKL--SSCDSFTSTINELNHCLSLRTYLVGNSLSLADLCVWATLK--GNAAWQEQL 138

Qy        293 --VTVPANVQRWMRSCENLAPFNT 314
                | :|:||    |    | :
Db        139 KQKKAPVHVKRWFGFLEAQQAFQS 162


RESULT 4
US-08-756-317-10
; Sequence 10, Application US/08756317
; Patent No. 5849894
;  GENERAL INFORMATION:
;    APPLICANT:  Clemente, Thomas E.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Mitsky, Timothy A.
;    APPLICANT:  Stark, David M.
;    TITLE OF INVENTION:  Improved Rhodospirillum Rubrum
;    TITLE OF INVENTION:  Poly-B-Hydroxyalkonoate Synthase
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  TX
;      COUNTRY:  USA
;      ZIP:  77210-4433
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:57:26 ; Search time 50.2443 Seconds
(without alignments)
2084.158 Million cell updates/sec

Title: US-10-622-817-2
Perfect score: 1667
Sequence: 1 MPMYQVKPYHGGGAPLRVEL.....RWMRSCENLAPFNTALKLLK 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1380268 seqs, 327241040 residues

Total number of hits satisfying chosen parameters: 1380268

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1630	97.8	312	16	US-10-463-676-4	Sequence 4, Appli
2	1191	71.4	229	16	US-10-463-676-6	Sequence 6, Appli
3	832	49.9	161	16	US-10-463-676-5	Sequence 5, Appli
4	102	6.1	201	15	US-10-369-493-13717	Sequence 13717, A
5	101	6.1	826	10	US-09-746-660A-92	Sequence 92, Appl
6	101	6.1	833	10	US-09-746-660A-90	Sequence 90, Appl
7	101	6.1	1221	9	US-09-919-891-2	Sequence 2, Appli
8	101	6.1	1221	9	US-09-738-626-5163	Sequence 5163, Ap
9	101	6.1	1221	15	US-10-450-055-2	Sequence 2, Appli
10	101	6.1	1221	17	US-10-494-675-156	Sequence 156, App
11	100.5	6.0	687	15	US-10-425-114-58897	Sequence 58897, A
12	100.5	6.0	1200	15	US-10-282-122A-53993	Sequence 53993, A
13	97.5	5.8	652	15	US-10-264-049-3001	Sequence 3001, Ap
14	97.5	5.8	711	16	US-10-437-963-175483	Sequence 175483,
15	97	5.8	559	9	US-09-364-847-21	Sequence 21, Appl
16	97	5.8	856	9	US-09-364-847-33	Sequence 33, Appl
17	97	5.8	856	9	US-09-364-847-35	Sequence 35, Appl
18	97	5.8	5215	9	US-09-861-289-2	Sequence 2, Appli
19	97	5.8	5215	9	US-09-860-846-2	Sequence 2, Appli
20	97	5.8	5215	10	US-09-988-384B-2	Sequence 2, Appli
21	97	5.8	5215	10	US-09-836-821-2	Sequence 2, Appli
22	97	5.8	5215	14	US-10-271-889-45	Sequence 45, Appl
23	96.5	5.8	914	16	US-10-437-963-153870	Sequence 153870,
24	96.5	5.8	1148	16	US-10-655-799-29	Sequence 29, Appl
25	96	5.8	606	16	US-10-437-963-115189	Sequence 115189,
26	95.5	5.7	455	15	US-10-425-114-42633	Sequence 42633, A
27	95	5.7	559	9	US-09-821-016-1	Sequence 1, Appli
28	95	5.7	559	9	US-09-820-952A-1	Sequence 1, Appli
29	95	5.7	559	14	US-10-218-519-1	Sequence 1, Appli
30	95	5.7	559	14	US-10-259-632-1	Sequence 1, Appli
31	95	5.7	559	14	US-10-266-787-1	Sequence 1, Appli
32	95	5.7	559	14	US-10-252-518-1	Sequence 1, Appli
33	95	5.7	559	17	US-10-914-244-1	Sequence 1, Appli
34	95	5.7	722	15	US-10-369-493-10013	Sequence 10013, A
35	95	5.7	936	14	US-10-032-585-7442	Sequence 7442, Ap
36	94.5	5.7	955	15	US-10-282-122A-69912	Sequence 69912, A
37	94	5.6	203	15	US-10-369-493-8604	Sequence 8604, Ap
38	94	5.6	468	15	US-10-424-599-260388	Sequence 260388,
39	94	5.6	1440	16	US-10-437-963-136277	Sequence 136277,
40	93.5	5.6	569	16	US-10-437-963-123904	Sequence 123904,
41	93.5	5.6	733	15	US-10-424-599-243766	Sequence 243766,
42	92	5.5	739	15	US-10-424-599-251522	Sequence 251522,
43	92	5.5	1206	15	US-10-282-122A-63838	Sequence 63838, A
44	91.5	5.5	1149	16	US-10-655-799-30	Sequence 30, Appl
45	91.5	5.5	1192	15	US-10-282-122A-64644	Sequence 64644, A

ALIGNMENTS

RESULT 1
US-10-463-676-4
; Sequence 4, Application US/10463676
; Publication No. US20040175375A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Sunghoon
; APPLICANT: Park, Bum-Joon
; TITLE OF INVENTION: Method for Treating Cancer Using P38/JTV-1 and Method
; TITLE OF INVENTION: for Screening Pharmaceutical Composition for Treating Cancer
; FILE REFERENCE: 012679-091
; CURRENT APPLICATION NUMBER: US/10/463,676
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: KR 10-2003-13058
; PRIOR FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)...(312)
; OTHER INFORMATION: 1-312 amino acid sequence of p38/JTV-1
US-10-463-676-4

Query Match 97.8%; Score 1630; DB 16; Length 312;
Best Local Similarity 100.0%; Pred. No. 6.4e-153;
Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWMRSCENLAPF 312
              ||||||||||||
Db        301 RWMRSCENLAPF 312


RESULT 2
US-10-463-676-6
; Sequence 6, Application US/10463676
; Publication No. US20040175375A1
; GENERAL INFORMATION:
;  APPLICANT: Kim, Sunghoon
;  APPLICANT:  Park, Bum-Joon
;  TITLE OF INVENTION: Method for Treating Cancer Using P38/JTV-1 and Method
;  TITLE OF INVENTION:  for Screening Pharmaceutical Composition for Treating Cancer
;  FILE REFERENCE: 012679-091
;  CURRENT APPLICATION NUMBER: US/10/463,676
;  CURRENT FILING DATE:  2003-06-18
;  PRIOR APPLICATION NUMBER: KR 10-2003-13058
;  PRIOR FILING DATE: 2003-03-03
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 229
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)...(229)
;   OTHER INFORMATION: 84-312 amino acid sequence p38/JTV-1
US-10-463-676-6

  Query Match             71.4%;  Score 1191;  DB 16;  Length 229;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-109;
  Matches  229;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         84 DADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLC 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 DADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLC 60

Qy        144 EHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQDYQLGFTLIWKNVPKTQMKFSIQT 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQDYQLGFTLIWKNVPKTQMKFSIQT 120

Qy        204 MCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSAL 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 MCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSAL 180

Qy        264 GKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQRWMRSCENLAPF 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQRWMRSCENLAPF 229


RESULT 3
US-10-463-676-5
; Sequence 5, Application US/10463676
; Publication No. US20040175375A1
; GENERAL INFORMATION:
;  APPLICANT: Kim, Sunghoon
;  APPLICANT:  Park, Bum-Joon
;  TITLE OF INVENTION: Method for Treating Cancer Using P38/JTV-1 and Method
;  TITLE OF INVENTION:  for Screening Pharmaceutical Composition for Treating Cancer
;  FILE REFERENCE: 012679-091
;  CURRENT APPLICATION NUMBER: US/10/463,676
;  CURRENT FILING DATE:  2003-06-18
;  PRIOR APPLICATION NUMBER: KR 10-2003-13058
;  PRIOR FILING DATE: 2003-03-03
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 161
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)...(161)
;   OTHER INFORMATION: 1-161 amino acid sequence of p38/JTV-1
US-10-463-676-5

  Query Match             49.9%;  Score 832;  DB 16;  Length 161;
  Best Local Similarity   100.0%;  Pred. No. 3.9e-74;
  Matches  161;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSV 161
              |||||||||||||||||||||||||||||||||||||||||
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSV 161


RESULT 4
US-10-369-493-13717
; Sequence 13717, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 13717
;   LENGTH: 201
;   TYPE: PRT
;   ORGANISM: Pseudomonas fluorescens
US-10-369-493-13717

  Query Match              6.1%;  Score 102;  DB 15;  Length 201;
  Best Local Similarity   30.5%;  Pred. No. 0.17;
  Matches   32;  Conservative   15;  Mismatches   36;  Indels   22;  Gaps    3;

Qy        211 GNIARFLFSLFGQKHNAVNA-TLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSALGKSPWL 269
              |  |  | ::|| | ||    |   ::: :          |:  |||:|:|
Db        106 GPAAARLITVFGAKFNAEEVITRAHNFLKV----------MDLELGKTPYL 146

Qy        270 AGNELTVADVVLWSVLQQI--GGCSVTVPANVQRWMRSCENLAPF 312
              || | |:|||  :| :     |  |:   |||: |:   | |  |
Db        147 AGTEPTIADVSAYSYIAHAPEGNVSLDDYANVRAWLSRIEALPGF 191
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:34:50 ; Search time 13.3213 Seconds
(without alignments)
2311.294 Million cell updates/sec

Title: US-10-622-817-2
Perfect score: 1667
Sequence: 1 MPMYQVKPYHGGGAPLRVEL.....RWMRSCENLAPFNTALKLLK 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	125.5	7.5	719	2	T52043	probable glutamate
2	111	6.7	728	2	T01200	probable glutamate
3	105.5	6.3	222	2	G82441	probable glutathio
4	97.5	5.8	218	2	A29036	glutathione transf
5	97.5	5.8	645	2	JC5517	Cu/RNA helicase II
6	97	5.8	559	2	A38604	poly(3-hydroxyalka
7	96.5	5.8	174	2	S74241	aminoacyl-tRNA syn
8	96	5.8	1675	2	S70770	hypothetical prote
9	94	5.6	218	2	B29231	glutathione transf
10	93	5.6	393	2	D64391	flavoprotein - Met
11	92	5.5	210	2	B85594	probable transfera
12	92	5.5	210	2	F64821	probable glutathio
13	92	5.5	210	2	F90743	probable transfera
14	92	5.5	744	2	AE2703	hypothetical prote
15	92	5.5	764	2	E97485	(P)PPGPP synthetas
16	92	5.5	867	2	S72842	methionine synthas
17	92	5.5	1036	2	AF2338	two-component sens
18	92	5.5	1206	2	E87072	hypothetical prote
19	92	5.5	2954	2	T14156	kinesin-related pr
20	91.5	5.5	218	1	XURTG4	glutathione transf
21	91.5	5.5	1192	2	G70513	5-methyltetrahydro
22	91	5.5	2279	2	T42531	acetyl-CoA carboxy
23	91	5.5	2280	2	T38906	acetyl-CoA carboxy
24	90.5	5.4	359	2	T37921	ras-associated pro
25	90	5.4	710	2	S67098	probable membrane
26	89.5	5.4	218	2	B34159	glutathione transf
27	89.5	5.4	1365	1	BVBYK5	killer toxin resis
28	89	5.3	2480	2	D84904	hypothetical prote
29	88.5	5.3	209	1	XUFF11	glutathione transf
30	88.5	5.3	1007	2	S48535	rho-type GTPase-ac
31	88	5.3	207	2	D72272	conserved hypothet
32	88	5.3	218	2	B28946	glutathione transf
33	88	5.3	437	2	I51238	translation elonga
34	87.5	5.2	214	2	B46681	glutathione transf
35	87.5	5.2	1065	2	A43421	SEC8 protein - yea
36	87.5	5.2	1292	2	T31462	probable magnesium
37	87.5	5.2	1714	1	S18644	multifunctional am
38	87.5	5.2	2586	2	T21676	hypothetical prote
39	87	5.2	586	2	C71695	ctp synthase (pyrG
40	86.5	5.2	208	1	S43851	glutathione transf
41	86.5	5.2	379	2	D97198	methyl-accepting c
42	86.5	5.2	436	2	I51237	translation elonga
43	86.5	5.2	439	2	D86440	unknown protein [i
44	86.5	5.2	749	2	A57276	MAD1 protein - yea
45	86	5.2	377	2	F69172	hypothetical prote

ALIGNMENTS

RESULT 1
T52043
probable glutamate-tRNA ligase (EC 6.1.1.17) [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004
C;Accession: T52043
R;Day, I.S.; Golovkin, M.; Reddy, A.S.
Biochim. Biophys. Acta 1399, 219-24, 1998
A;Title: Cloning of the cDNA for glutamyl-tRNA synthetase from Arabidopsis thaliana.
A;Reference number: Z24836; MUID:9765600; PMID:9765600
A;Accession: T52043
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-719 <DAY>
A;Cross-references: UNIPROT:O82462; EMBL:AF067773; PIDN:AAC36469.1
C;Superfamily: yeast glutamate-tRNA ligase; glutamine-tRNA ligase homology
C;Keywords: ligase

Query Match 7.5%; Score 125.5; DB 2; Length 719;
Best Local Similarity 27.3%; Pred. No. 0.017;
Matches 48; Conservative 26; Mismatches 53; Indels 49; Gaps 9;

```
Qy        128 PASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQDYQLGFTL 187
              |  |||||::|   |          |: : ::  :||  :   | :  |        | |:
Db         10 PESPPLSVIVALSLSASPV----TIDSSAAATTVPSFV---FSDGRKLN-------GATV 55

Qy        188 IWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA-IFQLKE 246
              : : |                  |  |: |   :|  :|| ::: || ||| | :|
Db         56 LLRYV-----------------GRSAKKLPDFYG--NNAFDSSQIDEWVDYASVF--SS 93

Qy        247 GSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQRW 302
              ||  | |     ::  |  | :| |: |::||| :|| |   |          |||
Db         94 GSEFENAC--GRVDKYLESSTFLVGHSLSIADVAIWSALAGTG----------QRW 137
```

RESULT 2
T01200
probable glutamate-tRNA ligase (EC 6.1.1.17) F21E10.12 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: T01200
R;Davidson, S.; Rohlfing, T.; David, M.; O'Brian, D.
submitted to the EMBL Data Library, April 1998
A;Description: The sequence of A. thaliana F21E10.
A;Reference number: Z14258
A;Accession: T01200
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-728 <DAV>
A;Cross-references: UNIPROT:O65253; EMBL:AF058914; NID:g3047074; PID:g3047084; GSPDB:GN0
A;Experimental source: cultivar Columbia
C;Genetics:

```
A;Gene: ATSP:F21E10.12
A;Map position: 5
A;Introns: 47/2; 89/3; 141/1; 503/3; 659/3
C;Superfamily: yeast glutamate-tRNA ligase; glutamine-tRNA ligase homology
C;Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis
F;223-499/Domain: glutamine-tRNA ligase homology <EGL>

  Query Match             6.7%;  Score 111;  DB 2;  Length 728;
  Best Local Similarity   25.9%;  Pred. No. 0.3;
  Matches   48;  Conservative   26;  Mismatches   53;  Indels   58;  Gaps   10;

Qy        128 PASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQDYQLGFTL 187
              | |||||::|   |          |: : ::  :||  :   | :  |         | |:
Db         10 PESPPLSVIVALSLSASPV----TIDSSAAATTVPSFV---FSDGRKLN-------GATV 55

Qy        188 IWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNAT---------LIDSWVD 238
              : : |                  |  |: |   :|  :|| :::           || |||
Db         56 LLRYV-----------------GRSAKKLPDFYG--NNAFDSSQVSILCINMKIDEWVD 95

Qy        239 IA-IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPA 297
               | :|    ||  | |     ::  |  | :| |: |::||| :|| |   |
Db         96 YASVF--SSGSSEFENAC--GRVDKYLESSTFLVGHSLSIADVAIWSALAGTG-------- 143

Qy        298 NVQRW 302
               |||
Db        144 --QRW 146


RESULT 3
G82441
probable glutathione S-transferase VCA0584 [imported] - Vibrio cholerae (strain N16961 s
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: G82441
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: G82441
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-222 <HEI>
A;Cross-references: UNIPROT:Q9KM05; GB:AE004389; GB:AE003853; NID:g9657989; PIDN:AAF9648
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VCA0584
A;Map position: 2
C;Superfamily: hypothetical protein b2302

  Query Match             6.3%;  Score 105.5;  DB 2;  Length 222;
  Best Local Similarity   25.7%;  Pred. No. 0.17;
  Matches   28;  Conservative   22;  Mismatches   38;  Indels   21;  Gaps    4;

Qy        197 MKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIAIFQLKEGSSKEKAAVF 256
              : | :  : |: |: |:    :  | :|       || :             ||   :|
Db         97 LMFQMSGVGPMMGQANV---FYRYFPEKIQPA----IDRY-----------QKEGRRLF 137

Qy        257 RSMNSALGKSPWLAGNELTVADVVL--WSVLQQIGGCSVTVPANVQRWM 303
               |:  | ::|:|||:| |:||:    |  : :  | |:    ::||||
Db        138 EVMDGQLAQNPYLAGDEYTIADIATFPWVRIHEWSGISIDGLTHLQRWM 186


RESULT 4
A29036
glutathione transferase (EC 2.5.1.18) Yb3 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C;Accession: A29036
R;Abramovitz, M.; Listowsky, I.
J. Biol. Chem. 262, 7770-7773, 1987
A;Title: Selective expression of a unique glutathione S-transferase Yb3 gene in rat brai
A;Reference number: A29036; MUID:87222405; PMID:3584141
A;Accession: A29036
A;Molecule type: mRNA
A;Residues: 1-218 <ABR>
A;Cross-references: UNIPROT:P08009; GB:J02744; NID:g204512; PIDN:AAA41292.1; PID:g20451
C;Superfamily: glutathione transferase
C;Keywords: transferase

  Query Match             5.8%;  Score 97.5;  DB 2;  Length 218;
  Best Local Similarity   25.6%;  Pred. No. 0.78;
  Matches   33;  Conservative   24;  Mismatches   41;  Indels   31;  Gaps    5;

Qy        179 QDYQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNAT------- 231
              : ::||  | : |:|     : |     |     | |:|    |:|||    |
Db         49 EKFKLG--LDFPNLP-----YLIDGSHKITQSNAILRYL----GRKHNLCGETEEERIRV 97

Qy        232 ------LIDSWVDIA-------IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVAD 278
                   |:|: : :|         :|| |  ::   : |  :  ||| || ||:::|  |
Db         98 DILENQLMDNRMVLARLCYNPDFEKLKPGYLEQLPGMMRLYSEFLGKRPWFAGDKITFVD 157

Qy        279 VVLWSVLQQ 287
               : : ||::
Db        158 FIAYDVLER 166


RESULT 5
JC5517
Gu/RNA helicase II binding protein - human
C;Species: Homo sapiens (man)
C;Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 05-Nov-1999
C;Accession: JC5517
R;Valdez, B.C.; Henning, D.; Perlaky, L.; Busch, R.K.; Busch, H.
Biochem. Biophys. Res. Commun. 234, 335-340, 1997
A;Title: Cloning and characterization of Gu/RH-II binding protein.
A;Reference number: JC5517; MUID:97320420; PMID:9177271
A;Accession: JC5517
A;Molecule type: mRNA
A;Residues: 1-645 <VAL>
A;Cross-references: GB:U78524; NID:g1696006; PIDN:AAB58488.1; PID:g1696007
A;Note: it is uncertain whether Met-4 or Met-6 is the initiator
C;Comment: This protein is localized to the nucleus and interacts with Gu/RNA helicase I
C;Keywords: phosphoprotein
F;50-58/Region: nuclear location signal
F;362-374/Region: nuclear location signal
F;514-517,551-554,592-595,606-609/Region: 4-residue repeats (N-T-S-L)
F;7,177,322,412,432,460,461,462,470,614,627/Binding site: phosphate (Ser) (covalent) #st
F;65,394,624/Binding site: phosphate (Thr) (covalent) #status predicted

  Query Match             5.8%;  Score 97.5;  DB 2;  Length 645;
  Best Local Similarity   20.5%;  Pred. No. 3.7;
  Matches   79;  Conservative   56;  Mismatches  129;  Indels  121;  Gaps   20;

Qy         16 LRV-ELPTCM-YRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQDDILKRLYELKAAV 73
              ||| ||   : |   | ||| :     | |: :       |:: :  :: :| :  |
Db          8 LRVSELQVLLGYAGRNKHGRKHELLTKALHLLKAG--CSPAVQMKIKELYRRRFPQK--- 62

Qy         74 DGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALKDIVINANPASPPL 133
                 | || ||| : |:  :  | ||: :                :  :  : :||| ||
Db         63 ------IMTP-ADLSIPNVHSSPMPATLSPS------------TIPQLTYDGHPASSPL 102

Qy        134 ---SLL-VLHRLLCEHF-RVLSTVHTHSSVKSVP-----ENLLK--CFGEQNKKQ----- 176
                 |||   | |   |   |  ||    :: :|     : |:|       | ::
Db        103 LPVSLLGPKHELELPHLTSALHPVHPDIKLQKLPFYDLLDELIKPTSLASDNSQRFRETC 162

Qy        177 ------PRQDYQLGFTLIWKNVPKTQMKFSIQ---------TMCPIEGEG---------- 211
                   |:|  |:  ::   ::  |:  |::|          | || |
Db        163 FAFALTPQQVQQISSSM---DISGTKCDFTVQVQLRFCLSETSCPQEDHFPPNLCVKVNT 219

Qy        212 ------------------------NIARFLFSLFGQKHNAVNATLIDSWV-------DI 239
                                     ||   :     :    |  |:: ||          :
Db        220 KPCSLPGYLPPTKNGVEPKRPSRPINITSLV-----RLSTTVPNTIVVSWTAEIGRNYSM 274

Qy        240 AIFQLKEGSS-------KEKAAVFRSMNSALGKSPWLAGNELTVADVVL-WSVLQQIGGC 291
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:33:04 ; Search time 60.8145 Seconds
(without alignments)
2694.512 Million cell updates/sec

Title: US-10-622-817-2
Perfect score: 1667
Sequence: 1 MPMYQVKPYHGGGAPLRVEL.....RWMRSCENLAPFNTALKLLK 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1667	100.0	320	1	MCA2_HUMAN	Q13155 homo sapien
2	1663	99.8	320	2	Q96CZ5	Q96cz5 homo sapien
3	1464	87.8	320	2	Q8R010	Q8r010 mus musculu
4	1464	87.8	320	2	Q8R2Y6	Q8r2y6 mus musculu
5	1444	86.6	320	1	MCA2_CRIGR	Q9wvm7 cricetulus
6	1260	75.6	280	2	Q8R3V2	Q8r3v2 mus musculu
7	1144.5	68.7	311	2	Q6DK86	Q6dk86 xenopus tro
8	1139.5	68.4	311	2	Q6INU4	Q6inu4 xenopus lae
9	1132.5	67.9	311	2	Q7ZYD7	Q7zyd7 xenopus lae
10	887.5	53.2	321	2	Q7T3C0	Q7t3c0 brachydanio
11	267.5	16.0	340	2	Q7Q7A3	Q7q7a3 anopheles g
12	224	13.4	301	2	Q7KUM5	Q7kum5 drosophila
13	224	13.4	313	2	Q8T060	Q8t060 drosophila
14	224	13.4	322	2	Q6NKM4	Q6nkm4 drosophila
15	214.5	12.9	334	1	MCA2_DROME	Q9vur3 drosophila
16	125.5	7.5	719	2	O82462	O82462 arabidopsis
17	111	6.7	728	2	O65253	O65253 arabidopsis
18	106.5	6.4	880	2	Q8ETT1	Q8ett1 oceanobacil
19	106.5	6.4	913	2	Q6DRB3	Q6drb3 brachydanio
20	106.5	6.4	913	2	Q6PFQ4	Q6pfq4 brachydanio
21	105.5	6.3	222	2	Q9KM05	Q9km05 vibrio chol
22	105	6.3	240	2	Q8BTR1	Q8btr1 mus musculu
23	105	6.3	1512	1	SYEP_MOUSE	Q8cgc7 mus musculu
24	105	6.3	1800	2	Q7PRA2	Q7pra2 anopheles g
25	104.5	6.3	922	2	Q6TEM5	Q6tem5 brachydanio
26	104	6.2	661	2	Q9VB85	Q9vb85 drosophila
27	104	6.2	702	2	Q8IHC6	Q8ihc6 drosophila
28	104	6.2	841	2	Q8IMQ2	Q8imq2 drosophila
29	103	6.2	328	2	Q86X73	Q86x73 homo sapien
30	103	6.2	869	2	Q6PD57	Q6pd57 homo sapien
31	103	6.2	1213	2	Q8FTD2	Q8ftd2 corynebacte
32	102.5	6.1	372	2	Q87XN7	Q87xn7 pseudomonas
33	102.5	6.1	715	2	Q9LIZ8	Q9liz8 oryza sativ
34	102	6.1	559	2	Q8KQ23	Q8kq23 pseudomonas
35	102	6.1	559	2	Q88D25	Q88d25 pseudomonas
36	102	6.1	1486	2	Q6TXE9	Q6txe9 rattus norv
37	101	6.1	1196	2	Q6M580	Q6m580 corynebacte
38	101	6.1	1221	2	Q8NQD1	Q8nqd1 corynebacte
39	100.5	6.0	454	2	Q6UDR4	Q6udr4 mus musculu
40	100.5	6.0	1201	2	Q6NH83	Q6nh83 corynebacte
41	99.5	6.0	651	1	PIA1_MOUSE	O88907 mus musculu
42	99	5.9	203	2	Q6W1I2	Q6w1i2 rhizobium s
43	98.5	5.9	222	2	Q8EGM4	Q8egm4 shewanella
44	98.5	5.9	582	2	Q66CB1	Q66cb1 yersinia ps
45	97.5	5.8	217	1	GTM3_RAT	P08009 rattus norv

ALIGNMENTS

RESULT 1
MCA2_HUMAN
ID MCA2_HUMAN STANDARD; PRT; 320 AA.
AC Q13155; Q9P1L2;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Multisynthetase complex auxiliary component p38 (JTV-1 protein)
DE (PRO0992).
GN Name=JTV1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96115582; PubMed=8666379;
RA Nicolaides N.C., Kinzler K.W., Vogelstein B.;
RT "Analysis of the 5' region of PMS2 reveals heterogeneous transcripts
RT and a novel overlapping gene.";
RL Genomics 29:329-334(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE OF 197-320 FROM N.A.
RC TISSUE=Fetal liver;
RA Zhang C., Yu Y., Zhang S., Wei H., Zhou G., Ouyang S., Luo L., Bi J.,
RA Liu M., He F.;
RT "Functional prediction of the coding sequences of 121 new genes
RT deduced by analysis of cDNA clones from human fetal liver.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
RN [4]

```
RP   INTERACTION WITH FUBP1.
RX   MEDLINE=22716800; PubMed=12819782; DOI=10.1038/ng1182;
RA   Kim M.J., Park B.-J., Kang Y.-S., Kim H.J., Park J.-H., Kang J.W.,
RA   Lee S.W., Han J.M., Lee H.-W., Kim S.;
RT   "Downregulation of FUSE-binding protein and c-myc by tRNA synthetase
RT   cofactor p38 is required for lung cell differentiation.";
RL   Nat. Genet. 34:330-336(2003).
CC   -!- FUNCTION: Probable core protein of the multisynthetase complex
CC       that serves as a template for the assembly of the supramolecular
CC       structure. Mediates ubiquitination of FUBP1 and its degradation by
CC       the proteasome.
CC   -!- SUBUNIT: Component of the multisynthetase complex which is
CC       comprised of a bifunctional glutamyl-prolyl-tRNA synthetase, the
CC       monospecific isoleucyl, leucyl, glutaminyl, methionyl, lysyl,
CC       arginyl, and aspartyl-tRNA synthetases as well as three auxiliary
CC       proteins, p18, p48 and p43. Binds FUBP1.
CC   -!- SIMILARITY: Contains 1 GST-like domain.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC       frameshift in position 312.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U24169; AAC50391.1; ALT_FRAME.
DR   EMBL; BC002853; AAH02853.1; -.
DR   EMBL; BC010156; AAH10156.1; -.
DR   EMBL; AF116615; AAF71039.1; -.
DR   H-InvDB; HIX0006460; -.
DR   MIM; 600859; -.
DR   InterPro; IPR010987; GST_C_like.
DR   InterPro; IPR004046; GST_Cterm.
DR   Pfam; PF00043; GST_C; 1.
KW   Protein biosynthesis.
SQ   SEQUENCE   320 AA;  35349 MW;  F253726B63C12BAB CRC64;

  Query Match           100.0%;  Score 1667;  DB 1;  Length 320;
  Best Local Similarity  100.0%;  Pred. No. 6e-127;
  Matches  320;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWMRSCENLAPFNTALKLLK 320
              ||||||||||||||||||||
Db        301 RWMRSCENLAPFNTALKLLK 320


RESULT 2
Q96CZ5
ID   Q96CZ5      PRELIMINARY;      PRT;   320 AA.
AC   Q96CZ5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   JTV1.
GN   Name=JTV1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RA   Strausberg R.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC013630; AAH13630.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   320 AA;  35335 MW;  19F14BF758612E08 CRC64;

  Query Match           99.8%;  Score 1663;  DB 2;  Length 320;
  Best Local Similarity  99.7%;  Pred. No. 1.3e-126;
  Matches  319;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIVINANPGSPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWMRSCENLAPFNTALKLLK 320
              ||||||||||||||||||||
Db        301 RWMRSCENLAPFNTALKLLK 320
```

```
RESULT 3
Q8R010
ID   Q8R010      PRELIMINARY;      PRT;   320 AA.
AC   Q8R010;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Jtv1-pending protein.
GN   Name=Jtv1;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N, and FVB/N-3; TISSUE=Mammary tumor;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RA   Strausberg R.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N-3; TISSUE=Mammary tumor;
RA   Strausberg R.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC024410; AAH24410.1; -.
DR   EMBL; BC026972; AAH26972.1; -.
DR   MGD; MGI:2385237; Jtv1.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   320 AA;  35396 MW;  1745D7EE4BC3670D CRC64;

  Query Match             87.8%;  Score 1464;  DB 2;  Length 320;
  Best Local Similarity   87.2%;  Pred. No. 1.8e-110;
  Matches  279;  Conservative   16;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||||||||||| |||||||||||||||||| ::  ||  ||||||  |  ||||||||||
Db          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              |||||||||||||||||||| ||||||||||:||||||| || ||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
              |||||||||||||||||||||| :||||||||||||:|||||:|||||| :|| | :
Db        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240
              ||||||||||||||||||||:|||||||||||||||||||||||||||  ||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300
              :|||:|||||||||||||||||:|||| |||||||||||||||| || |   | |||
Db        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300

Qy        301 RWMRSCENLAPFNTALKLLK 320
              ||::|||||||:|||:|||
Db        301 RWLKSCENLAPFSTALQLLK 320


RESULT 4
Q8R2Y6
ID   Q8R2Y6      PRELIMINARY;      PRT;   320 AA.
AC   Q8R2Y6;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Jtv1-pending protein.
GN   Name=Jtv1;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CZECH II; TISSUE=Mammary tumor;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CZECH II; TISSUE=Mammary tumor;
RA   Strausberg R.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC026958; AAH26958.1; -.
DR   MGD; MGI:2385237; Jtv1.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   320 AA;  35423 MW;  1C21F1A74C9882B4 CRC64;

  Query Match             87.8%;  Score 1464;  DB 2;  Length 320;
  Best Local Similarity   87.2%;  Pred. No. 1.8e-110;
  Matches  279;  Conservative   16;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||||||||||| |||||||||||||||||| ::  ||  ||||||  |  ||||||||||
Db          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTNPATDAGHVQETSEPSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              |||||||||||||||||||| ||||||||||:||||||| || ||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
```

```
Db        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Db        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Db        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300

Qy        301 RWMRSCENLAPFNTALKLLK 320

Db        301 RWLKSCENLAPFSTALQLLK 320


RESULT 5
MCA2_CRIGR
ID   MCA2_CRIGR     STANDARD;      PRT;   320 AA.
AC   Q9WVM7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Multisynthetase complex auxiliary component p38.
OS   Cricetulus griseus (Chinese hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Cricetulus.
OX   NCBI_TaxID=10029;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=99096915; PubMed=9878398; DOI=10.1006/jmbi.1998.2316;
RA   Quevillon S., Robinson J.-C., Berthonneau E., Siatecka M., Mirande M.;
RT   "Macromolecular assemblage of aminoacyl-tRNA synthetases:
RT   identification of protein-protein interactions and characterization of
RT   a core protein.";
RL   J. Mol. Biol. 285:183-195(1999).
CC   -!- FUNCTION: Probable core protein of the multisynthetase complex
CC       that serves as a template for the assembly of the supramolecular
CC       structure.
CC   -!- SUBUNIT: Component of the multisynthetase complex which is
CC       comprised of a bifunctional glutamyl-prolyl-tRNA synthetase, the
CC       monospecific isoleucyl, leucyl, glutaminyl, methionyl, lysyl,
CC       arginyl, and aspartyl-tRNA synthetases as well as three auxiliary
CC       proteins, p18, p48 and p43.
CC   -!- SIMILARITY: Contains 1 GST-like domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF072727; AAD38422.1; -.
DR   InterPro; IPR010987; GST_C_like.
DR   InterPro; IPR004046; GST_Cterm.
DR   Pfam; PF00043; GST_C; 1.
KW   Protein biosynthesis.
SQ   SEQUENCE   320 AA;  35433 MW;  6D24E033ABEC810A CRC64;

  Query Match           86.6%;  Score 1444;  DB 1;  Length 320;
  Best Local Similarity   86.6%;  Pred. No. 7.6e-109;
  Matches  277;  Conservative   16;  Mismatches   27;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Db          1 MPMYQVKSYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Db         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPVPLTTNALDLNLVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Db        121 DIVINANPASPPISLLVLHRLLCERYRVLSTVHTHSSVKNVPENLLKCFGEQARKQSRHE 180

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Db        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHSAVNLTLIDSWVDIA 240

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Db        241 MFQLKEGSSKERAAVFRSMNSALGKSPWLVGNELTVADVVLWSVLQQTGDGSGVAPANVQ 300

Qy        301 RWMRSCENLAPFNTALKLLK 320

Db        301 RWLKSCENLVPFSTALQLLK 320


RESULT 6
Q8R3V2
ID   Q8R3V2      PRELIMINARY;      PRT;   280 AA.
AC   Q8R3V2;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   JTV1.
GN   Name=Jtv1;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RA   Strausberg R.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC024480; AAH24480.1; -.
DR   MGD; MGI:2385237; Jtv1.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   280 AA;  31097 MW;  E5BE81498983FD2E CRC64;

  Query Match           75.6%;  Score 1260;  DB 2;  Length 280;
  Best Local Similarity   88.0%;  Pred. No. 5.5e-94;
  Matches  243;  Conservative   14;  Mismatches   19;  Indels    0;  Gaps    0;

Qy         45 QEESNLSLQALESRQDDILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTN 104
```

```
Db          5 QETSEPSLQALESRQDDILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATN 64

Qy        105 ALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPEN 164

Db         65 TLDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPEN 124

Qy        165 LLKCFGEQNKKQPRQDYQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQK 224

Db        125 LVKCFGEQARKQSRHEYQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQK 184

Qy        225 HNAVNATLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSV 284

Db        185 HNAVTLTLIDSWVDIAMFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSV 244

Qy        285 LQQIGGCSVTVPANVQRWMRSCENLAPFNTALKLLK 320

Db        245 LQQTGGSSGAAPTNVQRWLKSCENLAPFSTALQLLK 280


RESULT 7
Q6DK86
ID   Q6DK86      PRELIMINARY;      PRT;   311 AA.
AC   Q6DK86;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   MGC69221 protein.
GN   Name=MGC69221;
OS   Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8364;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Gerhard D.S.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC074561; AAH74561.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   311 AA;  34480 MW;  D98F27F73C466154 CRC64;

  Query Match           68.7%;  Score 1144.5;  DB 2;  Length 311;
  Best Local Similarity   68.8%;  Pred. No. 1.5e-84;
  Matches  220;  Conservative   43;  Mismatches   48;  Indels    9;  Gaps    3;

Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Db          1 MPMYRMQPYCGG--EIQVDLPTCMYRLPNVH------QAVSENLEQAADPAIQALESRQE 52

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Db         53 DILKRLYELKAAVDGLSKMIQTPDADLDVTDIIQTDEQSAPGPVVADLDSVLGKEYGALR 112

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Db        113 DIVINANPSLPPLSLLILHSLLCERYQVLSAVHTHSSIASIPEPLMKCFGDQLKMKPRHQ 172

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Db        173 YQLGFTLIWKDVPKPQMKFSIQNMCPIEGEGNIARFLFSLLGYKFNAINATLIDGWVDTA 232

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Db        233 IFQLREGSSKEKAAVLKAMNTALGKSPWLVGNELTVADIVSWCAVQQCGN-STAVPPSVQ 291

Qy        301 RWMRSCENLAPFNTALKLLK 320

Db        292 KWMKSCENLASFKSVLRFLK 311


RESULT 8
Q6INU4
ID   Q6INU4      PRELIMINARY;      PRT;   311 AA.
AC   Q6INU4;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   MGC80304 protein.
GN   Name=MGC80304;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
```

```
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RA   Klein S., Gerhard D.S.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC072178; AAH72178.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   311 AA;  34424 MW;  D5E8325C18D88751 CRC64;

  Query Match             68.4%;  Score 1139.5;  DB 2;  Length 311;
  Best Local Similarity   69.1%;  Pred. No. 3.8e-84;
  Matches  221;  Conservative   40;  Mismatches   50;  Indels    9;  Gaps    3;

Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||:::|| ||   ::|:||||||||||||  :      | : :: :: ::|||||||:
Db          1 MPMYRMQPYCGG--EIQVDLPTCMYRLPNVHRAT------AENHEQAADPAIQALESRQE 52

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||:|||||| :   :   ||: ||| |||||:
Db         53 DILKRLYELKAAVDGLSKMIQTPDADLDVTDIIQADEQSATGSVVADLDVVLGNDYGALR 112

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
              ||||||||:  |||||:|| |||| ::||| |||||||: |:|| |||||||:| | :|| :
Db        113 DIVINANPSLAPLSLLILHSLLCERYQVLSAVHTHSSITSIPEPLLKCFGDQLKMKPRHE 172

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240
              ||||||||||:||| ||||||| |||||||||| |||||| |   |||||||||| ||| |
Db        173 YQLGFTLIWKDVPKPQMKFSIQNMCPIEGEGNIGRFLFSLLGYTFNAVNATLIDGWVDTA 232

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300
              ||||:|||||||||| ::||:|||||||| |||||||||:| |  :|| |  |  || |||
Db        233 IFQLREGSSKEKAAVLKAMNTALGKSPWLVGNELTVADIVSWCAVQQCGN-STAVPPNVQ 291

Qy        301 RWMRSCENLAPFNTALKLLK 320
              :||:|||||| | : |: ||
Db        292 KWMKSCENLASFKSVLRFLK 311
```

```
RESULT 9
Q7ZYD7
ID   Q7ZYD7      PRELIMINARY;      PRT;   311 AA.
AC   Q7ZYD7;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Jtv1-prov protein.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC043832; AAH43832.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   311 AA;  34647 MW;  2F08C9DA60DD63BE CRC64;

  Query Match             67.9%;  Score 1132.5;  DB 2;  Length 311;
  Best Local Similarity   69.1%;  Pred. No. 1.4e-83;
  Matches  221;  Conservative   39;  Mismatches   51;  Indels    9;  Gaps    4;

Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||:::|| ||   ::|:|||||||| ||||      |   | :: :: ::|||||||:
Db          1 MPMYRMQPYCGG--EIQVDLPTCMYRLANVH-----QATSENH-EQAADPAIQALESRQE 52

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||:||| || :   :   ||: ||||:||||:
Db         53 DILKRLYELKAAVDGLSKMIQTPDADLDVTHIIQTDEHSAPGSVVADLDLVLGKEYGALR 112

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
              ||||||||: |||||:|| |||| ::||| ||||||: |:|| ||||||||: | : | :
Db        113 DIVINANPSLPPLSLLILHSLLCERYQVLSAVHTHSSITSIPEPLLKCFGEQLKMKLRHE 172

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240
              ||||||||||:||| ||||||| |||||||||| |||||||| | ||||||||| ||| |
Db        173 YQLGFTLIWKDVPKPQMKFSIQNMCPIEGEGNIGRFLFSLFGYKFNAVNATLIDGWVDTA 232

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300
              ||||::|||||||| ::||: ||||||| |||||||:| |  :|| |  |  || |||
Db        233 IFQLRDGSSKEKAAVLKAMNTTLGKSPWLVGNELTVADIVSWCAVQQCGN-STVVPPNVQ 291

Qy        301 RWMRSCENLAPFNTALKLLK 320
              :||:|||||| | : |: ||
Db        292 KWMKSCENLASFKSVLRFLK 311
```

```
RESULT 10
Q7T3C0
ID   Q7T3C0      PRELIMINARY;      PRT;   321 AA.
AC   Q7T3C0;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Zgc:63976.
GN   ORFNames=zgc:63976;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
```

```
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Strausberg R.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC053178; AAH53178.1; -.
DR   ZFIN; ZDB-GENE-040426-2652; zgc:63976.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   321 AA;  34852 MW;  BA8F6B951208244A CRC64;

  Query Match             53.2%;  Score 887.5;  DB 2;  Length 321;
  Best Local Similarity   55.6%;  Pred. No. 1.1e-63;
  Matches  184;  Conservative   50;  Mismatches   76;  Indels   21;  Gaps    8;

Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQE-ESNLSLQALESRQ 59
              ||||||||     | : |:||||||:||||| :  | : |   :|  | : :::||| ||
Db          1 MPMYQVKPV--SPADITVDLPTCMYKLPNVHAQ--GASLGEHALQNGEVDPTVKALEERQ 56

Qy         60 DDILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLT--------TNALDLNSV 111
              |:||::|||||| ||||:| : ||||||| : :        |||           ||: :
Db         57 DEILRKLYELKATVDGLAKTVTTPDADLDASTLAH-----TLTHTPADAVLRGTADLDDL 111

Qy        112 LGKDYGALKDIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGE 171
              |||| |||:|||||||||| |||||||| |||: |:|||:|| ||,| :||  || | |
Db        112 LGKDLGALRDIVINANPAQPPLSLLVLHALLCQRFQVLSSVHVHSSVSTVPAPLLSCLGP 171

Qy        172 QN-KKQPRQDYQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFG-QKHNAVN 229
              ::      |  :|||||||||:| | ||||| | ||||||||||:|||: | | :  : |:
Db        172 RHTHSYARHRFQLGFTLIWKDVSKLQMKFSTQNMCPIEGEGNVARFLYRLLGAEPRDPVS 231

Qy        230 ATLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIG 289
              |||:| ||| |:||| || |||:||| |::|:|||:|||| | | ::||:|    : | |
Db        232 ATLMDGWVDTALFQLAEGGSKERAAVLRALNAALGRSPWLLGQEFSLADIVSACCVLQTG 291

Qy        290 GCSVTVPANVQRWMRSCENLAPFNTALKLLK 320
                | : |||||||::||:||  |:    ||:
Db        292 QTS-SAPANVQRWLKSCQNLGYFSCVDPLLQ 321


RESULT 11
Q7Q7A3
ID   Q7Q7A3      PRELIMINARY;      PRT;   340 AA.
AC   Q7Q7A3;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   AgCP5808 (Fragment).
GN   Name=agCG50514; ORFNames=ENSANGG00000011827;
OS   Anopheles gambiae str. PEST.
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX   NCBI_TaxID=180454;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PEST;
RA   Anopheles Genome Sequencing Consortium;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AAAB01008960; EAA10870.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   340 AA;  37467 MW;  FD0D8525CD4EB976 CRC64;

  Query Match             16.0%;  Score 267.5;  DB 2;  Length 340;
  Best Local Similarity   29.2%;  Pred. No. 2.6e-13;
  Matches  100;  Conservative   47;  Mismatches  133;  Indels   63;  Gaps   13;

Qy          3 MYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGP-------APG---AGHV-------- 44
              ||::||       :  ||||||| |  |   |: |       | |   || |
Db         12 MYRLKPVM--NQDVVCELPTCMYTLKPVCAYSHDPKAFSADSATGLAAAGPVDPMSIEID 69

Qy         45 ---QEESNLSLQALESRQDDILKRLYELKAAVDGLSKMIQTPDADLDV-TNIIQADEPTT 100
                 |:  |  |: |  ||  :|::| |||       | |     :| :  |   | :|:|
Db         70 NLLQKAENDELRMLAERQQRVLQQLAELK-------KEIMAMRTELKLNANAPPAVQPST 122

Qy        101 -------LTTNALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCEHFRVLSTVH 153
                    |   ::|          |:| |:||:|   | ||| |  |  :   :
Db        123 PLKSKAQLKAEPINLT--------CLQDFVVNASPEYVPYSLLALKNLWKDRLNLQVECF 174

Qy        154 THSSVKSVPE------NLLKCFGEQNKKQPRQDYQLGFTLIWKNV-PKTQMKFSIQTMCP 206
              |||:|  : |      | :   |       ||    : |||||||   |:|  |  :  |
Db        175 THSTVPKLSEEALAFQNAVTASGTAAANLPR----IKVTLIWKNVGAYTEMITSPTSYVP 230

Qy        207 IEGEGNIARFLFSLFGQKHN---AVNATLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSAL 263
              | || || |:|      : |      |   :|| :|     | : : |::  : |:: : |
Db        231 ICGEVNILRYLSRCGPSEFNYEQQDNVDEVDSILDACYLLLNKQNVKQRQQILRTLGAKL 290

Qy        264 GKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQRWMRSC 306
              ||:    | :|:: |:   | ::|:     ||  :|   | |
Db        291 GKAAGFGGADLSLCDIAFTSAVKQV---QRTVAKDVNPNMHKC 330


RESULT 12
Q7KUM5
ID   Q7KUM5      PRELIMINARY;      PRT;   301 AA.
AC   Q7KUM5;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   CG12304-PB.
GN   ORFNames=CG12304;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA   Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
```

```
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA   Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426065; PubMed=12537568;
RA   Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA   Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA   George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA   Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA   Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA   Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT   "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT   melanogaster euchromatic genome sequence.";
RL   Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426070; PubMed=12537573;
RA   Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA   Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA   Ashburner M., Celniker S.E.;
RT   "The transposable elements of the Drosophila melanogaster euchromatin:
RT   a genomics perspective.";
RL   Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfied E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN   [5]
RP   SEQUENCE FROM N.A.
RG   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RG   FlyBase;
RL   Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003530; AAS64998.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   301 AA;  33241 MW;  35224E17FE03F3BE CRC64;

  Query Match             13.4%;  Score 224;  DB 2;  Length 301;
  Best Local Similarity   28.1%;  Pred. No. 7.6e-10;
  Matches   85;  Conservative   47;  Mismatches  120;  Indels   50;  Gaps   10;

Qy        18 VELPTCMYRLPNV--------HGRSYGPAPGAGHVQEESNLSLQALESRQDDILKRLYEL 69
             ::|||||| | ||          | |   :  |      :   |: |   ||: :||:| ||
Db        13 IKLPTCMYPLKNVSLAADSLASGSSTSASTSASTSSCDDTASVAA---RQEKVLKQLEEL 69

Qy        70 KAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALKDIVINANPA 129
             || :  :         | | |    :  : ||   |         |     |:|:||| :|
Db        70 KAQLGQIR-------AGLGVCG--KTFQHTTAFQNG-------GLKEVPLQDVVINGHPN 113

Qy       130 SPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSV-------PENLLKCFGEQNKKQPRQDYQ 182
             | :|| |       : :     |||::  :         || |     |   |:
Db       114 FIPYALLALKNAWRNLYTIDVKTFTHSTMADIGPAAREFEANLAKV--PVNPALPK---- 167

Qy       183 LGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATL---IDSWVDI 239
             :  ||||||   |:|  |      || || || |:|  :   ::    : |   ||  :||
Db       168 ISVTLIWKNCEHTEMISSPTMYVPIYGEVNIIRYLGRVGPAEYRYEGSPLCNEIDLVLDI 227

Qy       240 AIFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANV 299
             |:  : | : |: | ::  | |  :  |::::|||| ::| |       : :||
Db       228 CYQLLRCNTHKTQVAMVRLLDKRLQKQQYFGGSQMSVADVGVYSSL-------IRMPAVT 280

Qy       300 QR 301
             ::
Db       281 EK 282


RESULT 13
Q8T060
ID   Q8T060      PRELIMINARY;      PRT;   313 AA.
AC   Q8T060;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   LD25772p (Fragment).
GN   ORFNames=CG12304;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA   George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA   Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA   Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA   Celniker S.;
RL   Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY069537; AAL39682.2; -.
DR   FlyBase; FBgn0036515; CG12304.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   313 AA;  34522 MW;  5AF9C801877F8531 CRC64;

  Query Match             13.4%;  Score 224;  DB 2;  Length 313;
  Best Local Similarity   28.1%;  Pred. No. 8e-10;
  Matches   85;  Conservative   47;  Mismatches  120;  Indels   50;  Gaps   10;

Qy        18 VELPTCMYRLPNV--------HGRSYGPAPGAGHVQEESNLSLQALESRQDDILKRLYEL 69
             ::|||||| | ||          | |   :  |      :   |: |   ||: :||:| ||
Db        25 IKLPTCMYPLKNVSLAADSLASGSSTSASTSASTSSCDDTASVAA---RQEKVLKQLEEL 81

Qy        70 KAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALKDIVINANPA 129
             || :  :         | | |    :  : ||   |         |     |:|:||| :|
Db        82 KAQLGQIR-------AGLGVCG--KTFQHTTAFQNG-------GLKEVPLQDVVINGHPN 125

Qy       130 SPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSV-------PENLLKCFGEQNKKQPRQDYQ 182
             | :|| |       : :     |||::  :         || |     |   |:
Db       126 FIPYALLALKNAWRNLYTIDVKTFTHSTMADIGPAAREFEANLAKV--PVNPALPK---- 179
```

```
Qy        183 LGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATL---IDSWVDI 239
              :  ||||||   |:|  |      || || || |:|  :   ::    : |   ||  :||
Db        180 ISVTLIWKNCEHTEMISSPTMYVPIYGEVNIIRYLGRVGPAEYRYEGSPLCNEIDLVLDI 239

Qy        240 AIFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANV 299
                |:  : | : |: | ::  | |  :  |::::|||| ::| |       : :||
Db        240 CYQLLRCNTHKTQVAMVRLLDKRLQKQQYFGGSQMSVADVGVYSSL-------IRMPAVT 292

Qy        300 QR 301
              ::
Db        293 EK 294


RESULT 14
Q6NKM4
ID   Q6NKM4      PRELIMINARY;      PRT;   322 AA.
AC   Q6NKM4;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   LP12114p (Fragment).
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Carlson J., Chavez C., Frise E., George R., Pacleb J.,
RA   Park S., Wan K., Yu C., Rubin G.M., Celniker S.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BT012670; AAT08476.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   322 AA;  35528 MW;  82A15927DF928DAE CRC64;

  Query Match             13.4%;  Score 224;  DB 2;  Length 322;
  Best Local Similarity   28.1%;  Pred. No. 8.3e-10;
  Matches   85;  Conservative   47;  Mismatches  120;  Indels   50;  Gaps   10;

Qy         18 VELPTCMYRLPNV--------HGRSYGPAPGAGHVQEESNLSLQALESRQDDILKRLYEL 69
              ::|||||| | ||         | |   :  |     :    |: |   ||: :||:| ||
Db         34 IKLPTCMYPLKNVSLAADSLASGSSTSASTSASTSSCDDTASVAA---RQEKVLKQLEEL 90

Qy         70 KAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALKDIVINANPA 129
              || :  :        | | |    :  : ||   |         |      |:|:||| :|
Db         91 KAQLGQIR-------AGLGVCG--KTFQHTTAFQNG-------GLKEVPLQDVVINGHPN 134

Qy        130 SPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSV-------PENLLKCFGEQNKKQPRQDYQ 182
               | :|| |       : :     |||::  :          || |      |   |:
Db        135 FIPYALLALKNAWRNLYTIDVKTFTHSTMADIGPAAREFEANLAKV--PVNPALPK---- 188

Qy        183 LGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATL---IDSWVDI 239
              :  ||||||   |:|  |      || || || |:|  :   ::    : |   ||  :||
Db        189 ISVTLIWKNCEHTEMISSPTMYVPIYGEVNIIRYLGRVGPAEYRYEGSPLCNEIDLVLDI 248

Qy        240 AIFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANV 299
                |:  : | : |: | ::  | |  :  |::::|||| ::| |       : :||
Db        249 CYQLLRCNTHKTQVAMVRLLDKRLQKQQYFGGSQMSVADVGVYSSL-------IRMPAVT 301

Qy        300 QR 301
              ::
Db        302 EK 303


RESULT 15
MCA2_DROME
ID   MCA2_DROME     STANDARD;      PRT;   334 AA.
AC   Q9VUR3;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   25-JAN-2005 (Rel. 46, Last annotation update)
DE   Probable multisynthetase complex auxiliary component p38.
GN   ORFNames=CG12304;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   GENOME REANNOTATION.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
CC   -!- FUNCTION: Probable core protein of the multisynthetase complex
CC       that serves as a template for the assembly of the supramolecular
CC       structure (By similarity).
CC   -!- SUBUNIT: Component of the multisynthetase complex which is
CC       comprised of a bifunctional glutamyl-prolyl-tRNA synthetase, the
CC       monospecific isoleucyl, leucyl, glutaminyl, methionyl, lysyl,
CC       arginyl, and aspartyl-tRNA synthetases as well as three auxiliary
CC       proteins, p18, p48 and p43 (By similarity).
CC   -!- SIMILARITY: Contains 1 GST-like domain.
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE003530; AAF49612.1; -.
DR   IntAct; Q9VUR3; -.
DR   FlyBase; FBgn0036515; CG12304.
DR   InterPro; IPR010987; GST_C_like.
DR   InterPro; IPR004046; GST_Cterm.
DR   Pfam; PF00043; GST_C; 1.
KW   Protein biosynthesis.
FT   DOMAIN      280    327       GST-like.
SQ   SEQUENCE   334 AA;  36933 MW;  B68FD70AE621990F CRC64;

  Query Match             12.9%;  Score 214.5;  DB 1;  Length 334;
  Best Local Similarity   25.6%;  Pred. No. 5.2e-09;
  Matches   85;  Conservative   50;  Mismatches  120;  Indels   77;  Gaps   10;

Qy         18 VELPTCMYRLPNV----HGRSYGPAPGAGHVQEESNLSLQA------------------ 54
              ::|||||| | ||        : | :   |       |:   |:|
Db         13 IKLPTCMYPLKNVSLAADSLASGSSTSASTSASTSSCKLEANRIDRTGRNAATCALDLDS 72

Qy         55 ---------------LESRQDDILKRLYELKAAVDGLSKMIQTPDADLLDVTNIIQADEPT 99
                           : :||: :||:| |||| :  :         | | |    :  : |
Db         73 LGRQIQRLLKDDTASVAARQEKVLKQLEELKAQLGQIR-------AGLGVCG--KTFQHT 123

Qy        100 TLTTNALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVK 159
              |   |         |      |:|:||| :|    | :|| |        : :     |||::
Db        124 TAFQNG-------GLKEVPLQDVVINGHPNFIPYALLALKNAWRNLYTIDVKTFTHSTMA 176

Qy        160 SV-------PENLLKCFGEQNKKQPRQDYQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGN 212
               :          || |      |   |:    :  ||||||   |:|  |      || || |
Db        177 DIGPAAREFEANLAKV--PVNPALPK----ISVTLIWKNCEHTEMISSPTMYVPIYGEVN 230

Qy        213 IARFLFSLFGQKHNAVNATL---IDSWVDIAIFQLKEGSSKEKAAVFRSMNSALGKSPWL 269
              | |:|  :   ::    : |   ||  :||    |:  : | : |: | ::  | |  :
Db        231 IIRYLGRVGPAEYRYEGSPLCNEIDLVLDICYQLLRCNTHKTQVAMVRLLDKRLQKQQYF 290

Qy        270 AGNELTVADVVLWSVLQQIGGCSVTVPANVQR 301
               |::::|||| ::| |       : :||  ::
Db        291 GGSQMSVADVGVYSSL-------IRMPAVTEK 315


Search completed: February 23, 2005, 13:57:11
Job time : 62.9811 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:25:39 ; Search time 70.371 Seconds
(without alignments)
1758.725 Million cell updates/sec

Title: US-10-622-817-3
Perfect score: 1667
Sequence: 1 MPMYQVKPYHGGGAPLRVEL.....RWMRSCENLAPFNTALKLLK 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1667	100.0	320	7	ADE58613	Ade58613 Human Pro
2	1667	100.0	320	8	ABM80843	Abm80843 Tumour-as
3	1630	97.8	312	2	AAW25776	Aaw25776 JTV1 prot
4	1630	97.8	312	8	ADR86551	Adr86551 1-312 ami
5	1378	82.7	272	7	ADC10204	Adc10204 Human NOV
6	1191	71.4	229	8	ADR86553	Adr86553 84-312 am
7	832	49.9	161	8	ADR86552	Adr86552 1-161 ami
8	247	14.8	51	5	ABP01502	Abp01502 Human ORF
9	214.5	12.9	334	4	ABB62468	Abb62468 Drosophil
10	181	10.9	263	4	ABG23964	Abg23964 Novel hum
11	125.5	7.5	716	3	AAG41411	Aag41411 Arabidops
12	125.5	7.5	719	3	AAG41410	Aag41410 Arabidops
13	125.5	7.5	748	3	AAG41409	Aag41409 Arabidops
14	104	6.2	639	4	ABB62218	Abb62218 Drosophil
15	103	6.2	1512	4	AAU04349	Aau04349 Mammalian
16	103	6.2	1512	4	AAM78732	Aam78732 Human pro
17	103	6.2	1550	4	AAM79716	Aam79716 Human pro
18	102	6.1	201	8	ADS24684	Ads24684 Bacterial
19	101	6.1	559	2	AAR10682	Aar10682 Polyhydro
20	101	6.1	826	4	AAB79740	Aab79740 Corynebac
21	101	6.1	826	4	AAB79989	Aab79989 Corynebac
22	101	6.1	826	4	AAU71908	Aau71908 C. glutam
23	101	6.1	833	4	AAB79739	Aab79739 Corynebac
24	101	6.1	833	4	AAB79988	Aab79988 Corynebac
25	101	6.1	833	4	AAU71907	Aau71907 C. glutam
26	101	6.1	1221	4	AAG91409	Aag91409 C glutami
27	101	6.1	1221	5	ABG80321	Abg80321 C. glutam
28	101	6.1	1221	5	AAE18908	Aae18908 Corynebac
29	101	6.1	1221	7	ADD13589	Add13589 C. glutam
30	100.5	6.0	1200	6	ABU26069	Abu26069 Protein e
31	99.5	6.0	650	2	AAY29614	Aay29614 Mouse PIA
32	99.5	6.0	651	2	AAY29621	Aay29621 Mouse PIA
33	97.5	5.8	217	7	ADE63092	Ade63092 Rat Prote
34	97.5	5.8	217	7	ADD48150	Add48150 Rat Prote
35	97.5	5.8	217	7	ADD48158	Add48158 Rat Prote
36	97.5	5.8	217	7	ADD48154	Add48154 Rat Prote
37	97.5	5.8	217	7	ADE63084	Ade63084 Rat Prote
38	97.5	5.8	217	7	ADE63088	Ade63088 Rat Prote
39	97.5	5.8	218	2	AAR20033	Aar20033 Rat brain
40	97.5	5.8	218	7	ADE56502	Ade56502 Rat Prote
41	97.5	5.8	652	5	ABP41869	Abp41869 Human ova
42	97	5.8	559	3	AAY77178	Aay77178 S. venezu
43	97	5.8	2221	4	ABB61339	Abb61339 Drosophil
44	97	5.8	5215	8	ADL91930	Adl91930 Streptomy
45	96.5	5.8	1049	6	ABM67090	Abm67090 Photorhab

ALIGNMENTS

RESULT 1
ADE58613
ID ADE58613 standard; protein; 320 AA.
XX
AC ADE58613;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein Q13155, SEQ ID NO 4489.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
DR GENBANK; Q13155.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal

CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 320 AA;

Query Match 100.0%; Score 1667; DB 7; Length 320;
Best Local Similarity 100.0%; Pred. No. 5.4e-167;
Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWMRSCENLAPFNTALKLLK 320
              ||||||||||||||||||||
Db        301 RWMRSCENLAPFNTALKLLK 320
```

RESULT 2
ABM80843
ID ABM80843 standard; protein; 320 AA.
XX
AC ABM80843;
XX
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) polypeptide PRO81501, SEQ:2179.
XX
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2004030615-A2.
XX
PD 15-APR-2004.
XX
PF 29-SEP-2003; 2003WO-US028547.
XX
PR 02-OCT-2002; 2002US-0414971P.
XX
PA (GETH) GENENTECH INC.
XX
PI Wu TD, Zhang Z, Zhou Y;
XX
DR WPI; 2004-347921/32.
DR N-PSDB; ACN38579.
XX
PT New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
PS Claim 12; SEQ ID NO 2179; 7273pp; English.
XX
CC The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT polypeptide of the invention
XX
SQ Sequence 320 AA;

Query Match 100.0%; Score 1667; DB 8; Length 320;
Best Local Similarity 100.0%; Pred. No. 5.4e-167;
Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWMRSCENLAPFNTALKLLK 320
              ||||||||||||||||||||
Db        301 RWMRSCENLAPFNTALKLLK 320
```

RESULT 3
AAW25776

ID AAW25776 standard; protein; 312 AA.
XX
AC AAW25776;
XX
DT 19-DEC-1997 (first entry)
XX
DE JTV1 protein.
XX
KW JTV1; hPMS2; probe; detection; chromosome 7; deletion;
KW mismatch repair gene; hereditary non-polyposis colorectal cancer;
KW homologous recombination.
XX
OS Homo sapiens.
XX
PN WO9708312-A1.
XX
PD 06-MAR-1997.
XX
PF 26-AUG-1996; 96WO-US013598.
XX
PR 24-AUG-1995; 95US-00518862.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Vogelstein B, Kinzler KW, Nicolaides NC;
XX
DR WPI; 1997-179269/16.
DR N-PSDB; AAT86182.
XX
PT Novel chromosome 7 gene, JTV1 - used for detecting chromosome 7
PT deletions, and PMS2 promoter activity.
XX
PS Claim 5; Fig 2; 31pp; English.
XX
CC This sequence is JTV1 protein and is encoded by DNA isolated from human
CC chromosome 7. The JTV1 coding sequence is located upstream from hPMS2.
CC JTV1 cDNA can be used as probes to detect chromosome 7 deletions
CC involving JTV1. Due to the overlapping promoter regions, deletions of
CC JTV1 would also affect PMS2 (a mismatch repair gene) expression, leading
CC to hereditary non-polyposis colorectal cancer. JTV1 can also be used to
CC assay activity or competence of the PMS2 promoter region, the presence of
CC JTV1 suggesting that the PMS2 promoter is intact. JTV1 sequences can also
CC be used to guide homologous recombination at the PMS2 lccus
XX
SQ Sequence 312 AA;

Query Match 97.8%; Score 1630; DB 2; Length 312;
Best Local Similarity 100.0%; Pred. No. 4.3e-163;
Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWMRSCENLAPF 312
              ||||||||||||
Db        301 RWMRSCENLAPF 312
```

RESULT 4
ADR86551
ID ADR86551 standard; protein; 312 AA.
XX
AC ADR86551;
XX
DT 18-NOV-2004 (first entry)
XX
DE 1-312 amino acid sequence of p38/JTV-1 protein.
XX
KW p38/JTV-1; Cytostatic; cancer; leukemia; anticancer.
XX
OS Homo sapiens.
XX
PN EP1454628-A2.
XX
PD 08-SEP-2004.
XX
PF 09-SEP-2003; 2003EP-00020344.
XX
PR 03-MAR-2003; 2003KR-00013058.
XX
PA (UYSE-) UNIV SEOUL NAT IND FOUND.
XX
PI Kim S, Park B;
XX
DR WPI; 2004-627822/61.
DR N-PSDB; ADR86548.
XX
PT New isolated p38/JTV-1 protein, useful as medicament for treating cancer
PT e.g., stomach, liver, blood, bone, pancreatic, skin, head or neck cancer
PT and cutaneous or intraocular melanoma, as well as for screening new
PT anticancer agents.
XX
PS Claim 5; SEQ ID NO 4; 47pp; English.
XX
CC The present invention relates to an isolated p38/JTV-1 protein for use as
CC medicament. The p38/JTV-1 protein or the pharmaceutical composition is
CC useful as medicament for treating breast cancer, large intestinal cancer,
CC lung cancer, small cell lung cancer, stomach cancer, liver cancer, blood
CC cancer, bone cancer, pancreatic cancer, skin cancer, head or neck cancer,
CC cutaneous or intraocular melanoma, uterine sarcoma, ovarian cancer,
CC rectal cancer, anal cancer, colon cancer, fallopian tube carcinoma,
CC endometrial carcinoma, cervical cancer, vulval cancer, vaginal carcinoma,
CC Hodgkin's disease, esophageal cancer, small intestine cancer, endocrine
CC cancer, thyroid cancer, parathyroid cancer, adrenal cancer, soft tissue
CC tumour, urethral cancer, penile cancer, prostate cancer, chronic or acute
CC leukemia, lymphocytic lymphoma, bladder cancer, kidney cancer, ureter
CC cancer, renal cell carcinoma, renal pelvic carcinoma, CNS tumour, primary
CC CNS lymphoma, bone marrow tumour, brain stem nerve gliomas, pituitary
CC adenoma, or their combination. The protein is useful as a target for
CC screening new anticancer agents. The present sequence represents the 1-
CC 312 amino acid sequence of p38/JTV-1 protein.
XX
SQ Sequence 312 AA;

Query Match 97.8%; Score 1630; DB 8; Length 312;
Best Local Similarity 100.0%; Pred. No. 4.3e-163;
Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
```

```
Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWMRSCENLAPF 312
              ||||||||||||
Db        301 RWMRSCENLAPF 312
```

RESULT 5
ADC10204
ID ADC10204 standard; protein; 272 AA.
XX
AC ADC10204;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human NOVX polypeptide SEQ ID NO: 226.
XX
KW cytostatic; antidiabetic; anorectic; cerebroprotective; neuroprotective;
KW antiinflammatory; gene therapy; antisense therapy; thyromimetic; NOVX;
KW pathology; cancer; diabetes; obesity; endocrine disorder; CNS disorder;
KW inflammatory disorder; chromosome mapping; tissue typing;
KW predictive medicine.
XX
OS Homo sapiens.
XX
PN WO2003000842-A2.
XX
PD 03-JAN-2003.
XX
PF 04-JUN-2002; 2002WO-US017443.
XX
PR 04-JUN-2001; 2001US-0295607P.
PR 04-JUN-2001; 2001US-0295661P.
PR 06-JUN-2001; 2001US-0296404P.
PR 06-JUN-2001; 2001US-0296418P.
PR 07-JUN-2001; 2001US-0296575P.
PR 11-JUN-2001; 2001US-0297414P.
PR 12-JUN-2001; 2001US-0295573P.
PR 12-JUN-2001; 2001US-0297567P.
PR 14-JUN-2001; 2001US-0298285P.
PR 15-JUN-2001; 2001US-0298528P.
PR 18-JUN-2001; 2001US-0299133P.
PR 19-JUN-2001; 2001US-0299230P.
PR 21-JUN-2001; 2001US-0299949P.
PR 22-JUN-2001; 2001US-0300177P.
PR 26-JUN-2001; 2001US-0300883P.
PR 28-JUN-2001; 2001US-0301530P.
PR 28-JUN-2001; 2001US-0301550P.
PR 03-JUL-2001; 2001US-0302951P.
PR 31-JUL-2001; 2001US-0308890P.
PR 14-SEP-2001; 2001US-0322297P.
PR 25-SEP-2001; 2001US-0324669P.
PR 03-DEC-2001; 2001US-0337477P.
PR 14-DEC-2001; 2001US-0341562P.
PR 21-FEB-2002; 2002US-0358656P.
PR 21-FEB-2002; 2002US-0359122P.
PR 22-FEB-2002; 2002US-0358978P.
PR 22-FEB-2002; 2002US-0359034P.
PR 22-FEB-2002; 2002US-0359035P.
PR 22-FEB-2002; 2002US-0359121P.
PR 27-FEB-2002; 2002US-0359964P.
PR 01-MAR-2002; 2002US-0360858P.
PR 12-MAR-2002; 2002US-0363430P.
PR 12-MAR-2002; 2002US-0363676P.
PR 10-APR-2002; 2002US-0371346P.
PR 10-MAY-2002; 2002US-0379444P.
PR 04-JUN-2002; 2002US-00379444.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Agee ML, Anderson DW, Berghs C, Casman SJ, Catterton E;
PI Dipippo VA, Edinger SR, Eisen A, Ellerman K, Gangolli EA;
PI Gerlach VL, Gorman L, Guo X, Herrmann JL, Hjalt T, Ji W, Kekuda R;
PI Khramtsov NV, Li L, Liu X, Malyankar UM, Miller CE, Millet I;
PI Ort T, Padigaru M, Patturajan M, Pena CEA, Rastelli L, Rieger DK;
PI Rothenberg ME, Shenoy SG, Shimkets RA, Smithson G, Spaderna SK;
PI Spytek KA, Stone DJ, Vernet CAM, Zhong H, Zhong M, Alsobrook JP;
PI Burgess CE, Lepley DM;
XX
DR WPI; 2003-210149/20.
DR N-PSDB; ADC10203.
XX
PT New isolated NOVX polypeptides and nucleic acid molecules useful for
PT treating, preventing and diagnosing pathological conditions with NOVX-
PT associated disorders, such as cancer, obesity, diabetes and inflammatory
PT or CNS diseases.
XX
PS Claim 1; SEQ ID NO 226; 772pp; English.
XX
CC The invention relates to novel isolated polypeptides, mature form of the
CC polypeptide, a sequence that is 95% identical to the polypeptide or the
CC polypeptide comprising one or more conservative substitutions. The NOVX
CC polypeptide is useful for treating or preventing a pathology associated
CC with the polypeptide e.g. disorders associated with aberrant expression
CC or activity of the polypeptide, such as cancer, diabetes, obesity, and
CC endocrine, CNS and inflammatory disorders. They can also be used in
CC various detection and screening assays, chromosome mapping, tissue typing
CC and predictive medicine. This sequence corresponds to one of the
CC polypeptides of the invention.
XX
SQ Sequence 272 AA;

```
  Query Match             82.7%;  Score 1378;  DB 7;  Length 272;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-136;
  Matches  268;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          45 QEESNLSLQALESRQDDILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTN 104
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           5 QEESNLSLQALESRQDDILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTN 64

Qy         105 ALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPEN 164
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          65 ALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPEN 124

Qy         165 LLKCFGEQNKKQPRQDYQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQK 224
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         125 LLKCFGEQNKKQPRQDYQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQK 184

Qy         225 HNAVNATLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSV 284
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         185 HNAVNATLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSV 244

Qy         285 LQQIGGCSVTVPANVQRWMRSCENLAPF 312
               ||||||||||||||||||||||||||||
Db         245 LQQIGGCSVTVPANVQRWMRSCENLAPF 272
```

RESULT 6
ADR86553
ID ADR86553 standard; protein; 229 AA.
XX
AC ADR86553;
XX
DT 18-NOV-2004 (first entry)
XX
DE 84-312 amino acid sequence of p38/JTV-1 protein.
XX
KW p38/JTV-1; Cytostatic; cancer; leukemia; anticancer.
XX

OS Homo sapiens.
XX
PN EP1454628-A2.
XX
PD 08-SEP-2004.
XX
PF 09-SEP-2003; 2003EP-00020344.
XX
PR 03-MAR-2003; 2003KR-00013058.
XX
PA (UYSE-) UNIV SEOUL NAT IND FOUND.
XX
PI Kim S, Park B;
XX
DR WPI; 2004-627822/61.
DR N-PSDB; ADR86550.
XX
PT New isolated p38/JTV-1 protein, useful as medicament for treating cancer
PT e.g., stomach, liver, blood, bone, pancreatic, skin, head or neck cancer
PT and cutaneous or intraocular melanoma, as well as for screening new
PT anticancer agents.
XX
PS Claim 5; SEQ ID NO 6; 47pp; English.
XX
CC The present invention relates to an isolated p38/JTV-1 protein for use as
CC medicament. The p38/JTV-1 protein or the pharmaceutical composition is
CC useful as medicament for treating breast cancer, large intestinal cancer,
CC lung cancer, small cell lung cancer, stomach cancer, liver cancer, blood
CC cancer, bone cancer, pancreatic cancer, skin cancer, head or neck cancer,
CC cutaneous or intraocular melanoma, uterine sarcoma, ovarian cancer,
CC rectal cancer, anal cancer, colon cancer, fallopian tube carcinoma,
CC endometrial carcinoma, cervical cancer, vulval cancer, vaginal carcinoma,
CC Hodgkin's disease, esophageal cancer, small intestine cancer, endocrine
CC cancer, thyroid cancer, parathyroid cancer, adrenal cancer, soft tissue
CC tumour, urethral cancer, penile cancer, prostate cancer, chronic or acute
CC leukemia, lymphocytic lymphoma, bladder cancer, kidney cancer, ureter
CC cancer, renal cell carcinoma, renal pelvic carcinoma, CNS tumour, primary
CC CNS lymphoma, bone marrow tumour, brain stem nerve gliomas, pituitary
CC adenoma, or their combination. The protein is useful as a target for
CC screening new anticancer agents. The present sequence represents the 84-
CC 312 amino acid sequence of p38/JTV-1 protein.
XX
SQ Sequence 229 AA;

Query Match 71.4%; Score 1191; DB 8; Length 229;
Best Local Similarity 100.0%; Pred. No. 7.2e-117;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         84 DADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLC 143
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 DADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLC 60

Qy        144 EHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQDYQLGFTLIWKNVPKTQMKFSIQT 203
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQDYQLGFTLIWKNVPKTQMKFSIQT 120

Qy        204 MCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSAL 263
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 MCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSAL 180

Qy        264 GKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQRWMRSCENLAPF 312
               |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQRWMRSCENLAPF 229
```

RESULT 7
ADR86552
ID ADR86552 standard; protein; 161 AA.
XX
AC ADR86552;
XX
DT 18-NOV-2004 (first entry)
XX
DE 1-161 amino acid sequence of p38/JTV-1 protein.
XX
KW p38/JTV-1; Cytostatic; cancer; leukemia; anticancer.
XX
OS Homo sapiens.
XX
PN EP1454628-A2.
XX
PD 08-SEP-2004.
XX
PF 09-SEP-2003; 2003EP-00020344.
XX
PR 03-MAR-2003; 2003KR-00013058.
XX
PA (UYSE-) UNIV SEOUL NAT IND FOUND.
XX
PI Kim S, Park B;
XX
DR WPI; 2004-627822/61.
DR N-PSDB; ADR86549.
XX
PT New isolated p38/JTV-1 protein, useful as medicament for treating cancer
PT e.g., stomach, liver, blood, bone, pancreatic, skin, head or neck cancer
PT and cutaneous or intraocular melanoma, as well as for screening new
PT anticancer agents.
XX
PS Claim 5; SEQ ID NO 5; 47pp; English.
XX
CC The present invention relates to an isolated p38/JTV-1 protein for use as
CC medicament. The p38/JTV-1 protein or the pharmaceutical composition is
CC useful as medicament for treating breast cancer, large intestinal cancer,
CC lung cancer, small cell lung cancer, stomach cancer, liver cancer, blood
CC cancer, bone cancer, pancreatic cancer, skin cancer, head or neck cancer,
CC cutaneous or intraocular melanoma, uterine sarcoma, ovarian cancer,
CC rectal cancer, anal cancer, colon cancer, fallopian tube carcinoma,
CC endometrial carcinoma, cervical cancer, vulval cancer, vaginal carcinoma,
CC Hodgkin's disease, esophageal cancer, small intestine cancer, endocrine
CC cancer, thyroid cancer, parathyroid cancer, adrenal cancer, soft tissue
CC tumour, urethral cancer, penile cancer, prostate cancer, chronic or acute
CC leukemia, lymphocytic lymphoma, bladder cancer, kidney cancer, ureter
CC cancer, renal cell carcinoma, renal pelvic carcinoma, CNS tumour, primary
CC CNS lymphoma, bone marrow tumour, brain stem nerve gliomas, pituitary
CC adenoma, or their combination. The protein is useful as a target for
CC screening new anticancer agents. The present sequence represents the 1-
CC 161 amino acid sequence of p38/JTV-1 protein.
XX
SQ Sequence 161 AA;

Query Match 49.9%; Score 832; DB 8; Length 161;
Best Local Similarity 100.0%; Pred. No. 3.9e-79;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSV 161
              |||||||||||||||||||||||||||||||||||||||||
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSV 161
```

RESULT 8
ABP01502
ID ABP01502 standard; protein; 51 AA.
XX
AC ABP01502;
XX
DT 24-JUN-2002 (first entry)
XX

DE Human ORFX protein sequence SEQ ID NO:2986.
XX
KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypothyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW myasthenia gravis.
XX
OS Homo sapiens.
XX
PN WO200192523-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US010836.
XX
PR 30-MAY-2000; 2000US-0206132P.
PR 29-AUG-2000; 2000US-0228716P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach MD;
XX
DR WPI; 2002-106308/14.
DR N-PSDB; ABN17254.
XX
PT Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders.
XX
PS Disclosure; SEQ ID NO 2986; 1037pp; English.
XX
CC The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC in the specification). ABN15762 to ABN27252 encode the human ORFX
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage. N.B. The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 51 AA;

Query Match 14.8%; Score 247; DB 5; Length 51;
Best Local Similarity 96.1%; Pred. No. 4.5e-18;
Matches 49; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy         74 DGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALKDIVI 124
              ||||||| ||||||||||:|||||||||||||||||||||||||||||||
Db          1 DGLSKMIHTPDADLDVTNILQADEPTTLTTNALDLNSVLGKDYGALKDIVI 51
```

RESULT 9
ABB62468

ID ABB62468 standard; protein; 334 AA.
XX
AC ABB62468;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 14196.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL06571.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 14196; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 334 AA;

Query Match 12.9%; Score 214.5; DB 4; Length 334;
Best Local Similarity 25.6%; Pred. No. 2.8e-13;
Matches 85; Conservative 50; Mismatches 120; Indels 77; Gaps 10;

```
Qy          18 VELPTCMYRLPNV----HGRSYGPAPGAGHVQEESNLSLQA------------------ 54
               ::|||||| | ||       : | :  |      |:  |:|
Db          13 IKLPTCMYPLKNVSLAADSLASGSSTSASTSASTSSCKLEANRIDRTGRNAATCALDLDS 72

Qy          55 --------------LESRQDDILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPT 99
                           : :||: :||:| |||| :  :        | | |    :  : |
Db          73 LGRQIQRLLKDDTASVAARQEKVLKQLEELKAQLGQIR-------AGLGVCG--KTFQHT 123

Qy         100 TLTTNALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVK 159
               | |       |     |:|:||| :|   | :|| |       : :     |||:: 
Db         124 TAFQNG-------GLKEVPLQDVVINGHPNFIPYALLALKNAWRNLYTIDVKTFTHSTMA 176

Qy         160 SV-------PENLLKCFGEQNKKQPRQDYQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGN 212
                :             || |     |   |:    :  ||||||   |:|  |     || || |
Db         177 DIGPAAREFEANLAKV--PVNPALPK----ISVTLIWKNCEHTEMISSPTMYVPIYGEVN 230

Qy         213 IARFLFSLFGQKHNAVNATL---IDSWVDIAIFQLKEGSSKEKAAVFRSMNSALGKSPWL 269
               | |:|  :   ::    : |   ||  :||     |:  : | : |: | ::  | |  :
Db         231 IIRYLGRVGPAEYRYEGSPLCNEIDLVLDICYQLLRCNTHKTQVAMVRLLDKRLQKQQYF 290

Qy         270 AGNELTVADVVLWSVLQQIGGCSVTVPANVQR 301
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:36:20 ; Search time 17.81 Seconds
(without alignments)
1341.256 Million cell updates/sec

Title: US-10-622-817-3
Perfect score: 1667
Sequence: 1 MPMYQVKPYHGGGAPLRVEL.....RWMRSCENLAPFNTALKLLK 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1667	100.0	341	4	US-09-949-016-11312	Sequence 11312, A
2	1630	97.8	312	2	US-08-518-862C-2	Sequence 2, Appli
3	103	6.2	1512	3	US-09-443-184-48	Sequence 48, Appl
4	97	5.8	559	2	US-08-756-317-10	Sequence 10, Appl
5	97	5.8	559	4	US-09-091-609-4	Sequence 4, Appli
6	97	5.8	5215	3	US-09-105-537-2	Sequence 2, Appli
7	95	5.7	559	4	US-09-821-016-1	Sequence 1, Appli
8	95	5.7	559	4	US-10-259-632-1	Sequence 1, Appli
9	95	5.7	559	4	US-10-266-787-1	Sequence 1, Appli
10	92	5.5	2954	4	US-09-150-867-1	Sequence 1, Appli
11	91	5.5	474	4	US-09-248-796A-20321	Sequence 20321, A
12	90.5	5.4	359	3	US-09-540-824-2	Sequence 2, Appli
13	90.5	5.4	597	1	US-08-399-696-102	Sequence 102, App
14	88.5	5.3	1007	4	US-09-538-092-736	Sequence 736, App
15	88	5.3	427	4	US-09-949-016-11178	Sequence 11178, A
16	88	5.3	724	4	US-09-900-920-62	Sequence 62, Appl
17	88	5.3	883	4	US-09-976-239-2	Sequence 2, Appli
18	88	5.3	883	4	US-10-289-779B-2	Sequence 2, Appli
19	88	5.3	914	4	US-09-976-239-4	Sequence 4, Appli
20	88	5.3	914	4	US-10-289-779B-4	Sequence 4, Appli
21	86.5	5.2	761	3	US-09-012-710-13	Sequence 13, Appl
22	86.5	5.2	761	3	US-09-556-273-13	Sequence 13, Appl
23	86	5.2	211	4	US-09-328-352-6834	Sequence 6834, Ap
24	86	5.2	351	4	US-08-178-257-6	Sequence 6, Appli
25	86	5.2	471	3	US-09-134-001C-4904	Sequence 4904, Ap
26	86	5.2	1365	3	US-09-376-330-18	Sequence 18, Appl
27	86	5.2	1365	6	5194600-4	Patent No. 5194600
28	86	5.2	1365	6	5194600-4	Patent No. 5194600
29	85.5	5.1	1440	3	US-09-357-251-37	Sequence 37, Appl
30	85	5.1	559	4	US-10-253-509-1	Sequence 1, Appli
31	85	5.1	1135	2	US-08-469-537A-97	Sequence 97, Appl
32	84.5	5.1	579	4	US-09-107-433-4488	Sequence 4488, Ap
33	84.5	5.1	829	1	US-07-670-611-2	Sequence 2, Appli
34	84.5	5.1	829	1	US-08-220-674-2	Sequence 2, Appli
35	84.5	5.1	829	1	US-08-445-186-2	Sequence 2, Appli
36	84.5	5.1	829	1	US-08-446-549-2	Sequence 2, Appli
37	84.5	5.1	829	2	US-08-446-550-2	Sequence 2, Appli
38	83.5	5.0	778	4	US-09-583-110-3930	Sequence 3930, Ap
39	83.5	5.0	786	4	US-09-107-433-3893	Sequence 3893, Ap
40	83	5.0	559	2	US-09-052-339-1	Sequence 1, Appli
41	83	5.0	559	3	US-09-385-742B-1	Sequence 1, Appli
42	83	5.0	559	4	US-09-989-786-1	Sequence 1, Appli
43	83	5.0	767	3	US-08-836-567-8	Sequence 8, Appli
44	83	5.0	767	4	US-09-606-304-8	Sequence 8, Appli
45	82.5	4.9	222	4	US-09-248-796A-18146	Sequence 18146, A

ALIGNMENTS

RESULT 1
US-09-949-016-11312
; Sequence 11312, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11312
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11312

Query Match 100.0%; Score 1667; DB 4; Length 341;
Best Local Similarity 100.0%; Pred. No. 2.4e-181;
Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 81

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 141

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 201

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 261

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 321

Qy        301 RWMRSCENLAPFNTALKLLK 320
```

```
                    ||||||||||||||||||||
Db        322 RWMRSCENLAPFNTALKLLK 341


RESULT 2
US-08-518-862C-2
; Sequence 2, Application US/08518862C
; Patent No. 5843757
;  GENERAL INFORMATION:
;    APPLICANT:  Vogelstein, Bert
;    APPLICANT:  Kinzler, Kenneth W.
;    APPLICANT:  Nicolaides, Nicholas C.
;    TITLE OF INVENTION:  Human JTV1 Gene Overlaps PMS2 Gene
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Witcoff, Ltd.
;      STREET:  1001 G Street, N.W.
;      CITY:  Washington, D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/518,862C
;      FILING DATE:  24-AUG-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kagan, Sarah A.
;      REGISTRATION NUMBER:  32,141
;      REFERENCE/DOCKET NUMBER:  01107.49697
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-508-9100
;      TELEFAX:  202-508-9299
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  312 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-518-862C-2

  Query Match             97.8%;  Score 1630;  DB 2;  Length 312;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-177;
  Matches  312;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWMRSCENLAPF 312
              ||||||||||||
Db        301 RWMRSCENLAPF 312


RESULT 3
US-09-443-184-48
; Sequence 48, Application US/09443184A
; Patent No. 6372431
; GENERAL INFORMATION:
;  APPLICANT: Cunningham, Mary Jane
;  APPLICANT: Zweiger, Gary
;  APPLICANT: Kaser, Matthew R.
;  APPLICANT: Panzer, Scott
;  APPLICANT: Seilhammer, Jeffrey J.
;  APPLICANT: Yue, Henry
;  APPLICANT: Baughn, Mariah
;  APPLICANT: Azimzai, Yalda
;  APPLICANT: Lal, Preeti
;  TITLE OF INVENTION: MAMMALIAN TOXICOLOGICAL RESPONSE MARKERS
;  FILE REFERENCE: PC-0007 US
;  CURRENT APPLICATION NUMBER: US/09/443,184A
;  CURRENT FILING DATE: 1999-11-19
;  NUMBER OF SEQ ID NOS: 138
;  SOFTWARE: PERL Program
; SEQ ID NO 48
;   LENGTH: 1512
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6372431 2302721CD1
US-09-443-184-48

  Query Match             6.2%;  Score 103;  DB 3;  Length 1512;
  Best Local Similarity   22.1%;  Pred. No. 0.12;
  Matches   45;  Conservative   30;  Mismatches   69;  Indels   60;  Gaps    9;

Qy        124 INANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQDYQL 183
              :  |   |||  |    |  || :   ::    ||:   ||:|
Db          6 LTVNSGDPPLGAL----LAVEHVKDDVSI----SVEEGKENILH---------------- 41

Qy        184 GFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFL------FSLFGQKHNAVNATLIDSWV 237
              : :||  |:           :| |:|        |:|   | :  | || |:
Db         42 ----VSENVIFTDV------------NSILRYLARVATTAGLYGS--NLMEHTEIDHWL 82

Qy        238 DIAIFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCS----- 292
              : :  :|   |     :   :|  |    :| || |::||: :|: |:  |  :
Db         83 EFSATKL--SSCDSFTSTINELNHCLSLRTYLVGNSLSLADLCVWATLK--GNAAWQEQL 138

Qy        293 --VTVPANVQRWMRSCENLAPFNT 314
                | :|:||    |    | :
Db        139 KQKKAPVHVKRWFGFLEAQQAFQS 162


RESULT 4
US-08-756-317-10
; Sequence 10, Application US/08756317
; Patent No. 5849894
;  GENERAL INFORMATION:
;    APPLICANT:  Clemente, Thomas E.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Mitsky, Timothy A.
;    APPLICANT:  Stark, David M.
;    TITLE OF INVENTION:  Improved Rhodospirillum Rubrum
;    TITLE OF INVENTION:  Poly-B-Hydroxyalkonoate Synthase
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  TX
;      COUNTRY:  USA
;      ZIP:  77210-4433
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:57:26 ; Search time 50.2443 Seconds
(without alignments)
2084.158 Million cell updates/sec

Title: US-10-622-817-3
Perfect score: 1667
Sequence: 1 MPMYQVKPYHGGGAPLRVEL.....RWMRSCENLAPFNTALKLLK 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1380268 seqs, 327241040 residues

Total number of hits satisfying chosen parameters: 1380268

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1630	97.8	312	16	US-10-463-676-4	Sequence 4, Appli
2	1191	71.4	229	16	US-10-463-676-6	Sequence 6, Appli
3	832	49.9	161	16	US-10-463-676-5	Sequence 5, Appli
4	102	6.1	201	15	US-10-369-493-13717	Sequence 13717, A
5	101	6.1	826	10	US-09-746-660A-92	Sequence 92, Appl
6	101	6.1	833	10	US-09-746-660A-90	Sequence 90, Appl
7	101	6.1	1221	9	US-09-919-891-2	Sequence 2, Appli
8	101	6.1	1221	9	US-09-738-626-5163	Sequence 5163, Ap
9	101	6.1	1221	15	US-10-450-055-2	Sequence 2, Appli
10	101	6.1	1221	17	US-10-494-675-156	Sequence 156, App
11	100.5	6.0	687	15	US-10-425-114-58897	Sequence 58897, A
12	100.5	6.0	1200	15	US-10-282-122A-53993	Sequence 53993, A
13	97.5	5.8	652	15	US-10-264-049-3001	Sequence 3001, Ap
14	97.5	5.8	711	16	US-10-437-963-175483	Sequence 175483,
15	97	5.8	559	9	US-09-364-847-21	Sequence 21, Appl
16	97	5.8	856	9	US-09-364-847-33	Sequence 33, Appl
17	97	5.8	856	9	US-09-364-847-35	Sequence 35, Appl
18	97	5.8	5215	9	US-09-861-289-2	Sequence 2, Appli
19	97	5.8	5215	9	US-09-860-846-2	Sequence 2, Appli
20	97	5.8	5215	10	US-09-988-384B-2	Sequence 2, Appli
21	97	5.8	5215	10	US-09-836-821-2	Sequence 2, Appli
22	97	5.8	5215	14	US-10-271-889-45	Sequence 45, Appl
23	96.5	5.8	914	16	US-10-437-963-153870	Sequence 153870,
24	96.5	5.8	1148	16	US-10-655-799-29	Sequence 29, Appl
25	96	5.8	606	16	US-10-437-963-115189	Sequence 115189,
26	95.5	5.7	455	15	US-10-425-114-42633	Sequence 42633, A
27	95	5.7	559	9	US-09-821-016-1	Sequence 1, Appli
28	95	5.7	559	9	US-09-820-952A-1	Sequence 1, Appli
29	95	5.7	559	14	US-10-218-519-1	Sequence 1, Appli
30	95	5.7	559	14	US-10-259-632-1	Sequence 1, Appli
31	95	5.7	559	14	US-10-266-787-1	Sequence 1, Appli
32	95	5.7	559	14	US-10-252-518-1	Sequence 1, Appli
33	95	5.7	559	17	US-10-914-244-1	Sequence 1, Appli
34	95	5.7	722	15	US-10-369-493-10013	Sequence 10013, A
35	95	5.7	936	14	US-10-032-585-7442	Sequence 7442, Ap
36	94.5	5.7	955	15	US-10-282-122A-69912	Sequence 69912, A
37	94	5.6	203	15	US-10-369-493-8604	Sequence 8604, Ap
38	94	5.6	468	15	US-10-424-599-260388	Sequence 260388,
39	94	5.6	1440	16	US-10-437-963-136277	Sequence 136277,
40	93.5	5.6	569	16	US-10-437-963-123904	Sequence 123904,
41	93.5	5.6	733	15	US-10-424-599-243766	Sequence 243766,
42	92	5.5	739	15	US-10-424-599-251522	Sequence 251522,
43	92	5.5	1206	15	US-10-282-122A-63838	Sequence 63838, A
44	91.5	5.5	1149	16	US-10-655-799-30	Sequence 30, Appl
45	91.5	5.5	1192	15	US-10-282-122A-64644	Sequence 64644, A

ALIGNMENTS

RESULT 1
US-10-463-676-4
; Sequence 4, Application US/10463676
; Publication No. US20040175375A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Sunghoon
; APPLICANT: Park, Bum-Joon
; TITLE OF INVENTION: Method for Treating Cancer Using P38/JTV-1 and Method
; TITLE OF INVENTION: for Screening Pharmaceutical Composition for Treating Cancer
; FILE REFERENCE: 012679-091
; CURRENT APPLICATION NUMBER: US/10/463,676
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: KR 10-2003-13058
; PRIOR FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)...(312)
; OTHER INFORMATION: 1-312 amino acid sequence of p38/JTV-1
US-10-463-676-4

Query Match 97.8%; Score 1630; DB 16; Length 312;
Best Local Similarity 100.0%; Pred. No. 6.4e-153;
Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWMRSCENLAPF 312
              ||||||||||||
Db        301 RWMRSCENLAPF 312


RESULT 2
US-10-463-676-6
; Sequence 6, Application US/10463676
; Publication No. US20040175375A1
; GENERAL INFORMATION:
;  APPLICANT: Kim, Sunghoon
;  APPLICANT:  Park, Bum-Joon
;  TITLE OF INVENTION: Method for Treating Cancer Using P38/JTV-1 and Method
;  TITLE OF INVENTION:  for Screening Pharmaceutical Composition for Treating Cancer
;  FILE REFERENCE: 012679-091
;  CURRENT APPLICATION NUMBER: US/10/463,676
;  CURRENT FILING DATE:  2003-06-18
;  PRIOR APPLICATION NUMBER: KR 10-2003-13058
;  PRIOR FILING DATE: 2003-03-03
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 229
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)...(229)
;   OTHER INFORMATION: 84-312 amino acid sequence p38/JTV-1
US-10-463-676-6

  Query Match             71.4%;  Score 1191;  DB 16;  Length 229;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-109;
  Matches  229;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         84 DADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLC 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 DADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLC 60

Qy        144 EHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQDYQLGFTLIWKNVPKTQMKFSIQT 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQDYQLGFTLIWKNVPKTQMKFSIQT 120

Qy        204 MCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSAL 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 MCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSAL 180

Qy        264 GKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQRWMRSCENLAPF 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQRWMRSCENLAPF 229


RESULT 3
US-10-463-676-5
; Sequence 5, Application US/10463676
; Publication No. US20040175375A1
; GENERAL INFORMATION:
;  APPLICANT: Kim, Sunghoon
;  APPLICANT:  Park, Bum-Joon
;  TITLE OF INVENTION: Method for Treating Cancer Using P38/JTV-1 and Method
;  TITLE OF INVENTION:  for Screening Pharmaceutical Composition for Treating Cancer
;  FILE REFERENCE: 012679-091
;  CURRENT APPLICATION NUMBER: US/10/463,676
;  CURRENT FILING DATE:  2003-06-18
;  PRIOR APPLICATION NUMBER: KR 10-2003-13058
;  PRIOR FILING DATE: 2003-03-03
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 161
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)...(161)
;   OTHER INFORMATION: 1-161 amino acid sequence of p38/JTV-1
US-10-463-676-5

  Query Match             49.9%;  Score 832;  DB 16;  Length 161;
  Best Local Similarity   100.0%;  Pred. No. 3.9e-74;
  Matches  161;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSV 161
              |||||||||||||||||||||||||||||||||||||||||
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSV 161


RESULT 4
US-10-369-493-13717
; Sequence 13717, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 13717
;   LENGTH: 201
;   TYPE: PRT
;   ORGANISM: Pseudomonas fluorescens
US-10-369-493-13717

  Query Match              6.1%;  Score 102;  DB 15;  Length 201;
  Best Local Similarity   30.5%;  Pred. No. 0.17;
  Matches   32;  Conservative   15;  Mismatches   36;  Indels   22;  Gaps    3;

Qy        211 GNIARFLFSLFGQKHNAVNA-TLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSALGKSPWL 269
              |  |  | ::|| | ||    |   ::: :                   |:  |||:|:|
Db        106 GPAAARLITVFGAKFNAEEVITRAHNFLKV------------------MDLELGKTPYL 146

Qy        270 AGNELTVADVVLWSVLQQI--GGCSVTVPANVQRWMRSCENLAPF 312
              || | |:||| :| :    |  |:   |||: |:   | |  |
Db        147 AGTEPTIADVSAYSYIAHAPEGNVSLDDYANVRAWLSRIEALPGF 191
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:34:50 ; Search time 13.3213 Seconds
(without alignments)
2311.294 Million cell updates/sec

Title: US-10-622-817-3
Perfect score: 1667
Sequence: 1 MPMYQVKPYHGGGAPLRVEL.....RWMRSCENLAPFNTALKLLK 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	125.5	7.5	719	2	T52043	probable glutamate
2	111	6.7	728	2	T01200	probable glutamate
3	105.5	6.3	222	2	G82441	probable glutathio
4	97.5	5.8	218	2	A29036	glutathione transf
5	97.5	5.8	645	2	JC5517	Gu/RNA helicase II
6	97	5.8	559	2	A38604	poly(3-hydroxyalka
7	96.5	5.8	174	2	S74241	aminoacyl-tRNA syn
8	96	5.8	1675	2	S70770	hypothetical prote
9	94	5.6	218	2	B29231	glutathione transf
10	93	5.6	393	2	D64391	flavoprotein - Met
11	92	5.5	210	2	B85594	probable transfera
12	92	5.5	210	2	F64821	probable glutathio
13	92	5.5	210	2	F90743	probable transfera
14	92	5.5	744	2	AE2703	hypothetical prote
15	92	5.5	764	2	E97485	(P)PPGPP synthetas
16	92	5.5	867	2	S72842	methionine synthas
17	92	5.5	1036	2	AF2338	two-component sens
18	92	5.5	1206	2	E87072	hypothetical prote
19	92	5.5	2954	2	T14156	kinesin-related pr
20	91.5	5.5	218	1	XURTG4	glutathione transf
21	91.5	5.5	1192	2	G70513	5-methyltetrahydro
22	91	5.5	2279	2	T42531	acetyl-CoA carboxy
23	91	5.5	2280	2	T38906	acetyl-CoA carboxy
24	90.5	5.4	359	2	T37921	ras-associated pro
25	90	5.4	710	2	S67098	probable membrane
26	89.5	5.4	218	2	B34159	glutathione transf
27	89.5	5.4	1365	1	BVBYK5	killer toxin resis
28	89	5.3	2480	2	D84904	hypothetical prote
29	88.5	5.3	209	1	XUFF11	glutathione transf
30	88.5	5.3	1007	2	S48535	rho-type GTPase-ac
31	88	5.3	207	2	D72272	conserved hypothet
32	88	5.3	218	2	B28946	glutathione transf
33	88	5.3	437	2	I51238	translation elonga
34	87.5	5.2	214	2	B46681	glutathione transf
35	87.5	5.2	1065	2	A43421	SEC8 protein - yea
36	87.5	5.2	1292	2	T31462	probable magnesium
37	87.5	5.2	1714	1	S18644	multifunctional am
38	87.5	5.2	2586	2	T21676	hypothetical prote
39	87	5.2	586	2	C71695	ctp synthase (pyrG
40	86.5	5.2	208	1	S43851	glutathione transf
41	86.5	5.2	379	2	D97198	methyl-accepting c
42	86.5	5.2	436	2	I51237	translation elonga
43	86.5	5.2	439	2	D86440	unknown protein [i
44	86.5	5.2	749	2	A57276	MAD1 protein - yea
45	86	5.2	377	2	F69172	hypothetical prote

ALIGNMENTS

RESULT 1
T52043
probable glutamate-tRNA ligase (EC 6.1.1.17) [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004
C;Accession: T52043
R;Day, I.S.; Golovkin, M.; Reddy, A.S.
Biochim. Biophys. Acta 1399, 219-24, 1998
A;Title: Cloning of the cDNA for glutamyl-tRNA synthetase from Arabidopsis thaliana.
A;Reference number: Z24836; MUID:9765600; PMID:9765600
A;Accession: T52043
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-719 <DAY>
A;Cross-references: UNIPROT:O82462; EMBL:AF067773; PIDN:AAC36469.1
C;Superfamily: yeast glutamate-tRNA ligase; glutamine-tRNA ligase homology
C;Keywords: ligase

Query Match 7.5%; Score 125.5; DB 2; Length 719;
Best Local Similarity 27.3%; Pred. No. 0.017;
Matches 48; Conservative 26; Mismatches 53; Indels 49; Gaps 9;

```
Qy        128 PASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQDYQLGFTL 187
              | |||||::|   |          |: : ::  :||  :   | :  |        | |:
Db         10 PESPPLSVIVALSLSASPV----TIDSSAAATTVPSFV---FSDGRKLN-------GATV 55

Qy        188 IWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA-IFQLKE 246
              : : |                  |  |: |   :|  :|| ::: || ||| | :|
Db         56 LLRYV-----------------GRSAKKLPDFYG--NNAFDSSQIDEWVDYASVF--SS 93

Qy        247 GSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQRW 302
              ||  | |     ::  |  | :| |: |::||| :|| |   |          |||
Db         94 GSEFENAC--GRVDKYLESSTFLVGHSLSIADVAIWSALAGTG----------QRW 137
```

RESULT 2
T01200
probable glutamate-tRNA ligase (EC 6.1.1.17) F21E10.12 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: T01200
R;Davidson, S.; Rohlfing, T.; David, M.; O'Brian, D.
submitted to the EMBL Data Library, April 1998
A;Description: The sequence of A. thaliana F21E10.
A;Reference number: Z14258
A;Accession: T01200
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-728 <DAV>
A;Cross-references: UNIPROT:O65253; EMBL:AF058914; NID:g3047074; PID:g3047084; GSPDB:GN0
A;Experimental source: cultivar Columbia
C;Genetics:

A;Gene: ATSP:F21E10.12
A;Map position: 5
A;Introns: 47/2; 89/3; 141/1; 503/3; 659/3
C;Superfamily: yeast glutamate-tRNA ligase; glutamine-tRNA ligase homology
C;Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis
F;223-499/Domain: glutamine-tRNA ligase homology <EGL>

Query Match 6.7%; Score 111; DB 2; Length 728;
Best Local Similarity 25.9%; Pred. No. 0.3;
Matches 48; Conservative 26; Mismatches 53; Indels 58; Gaps 10;

```
Qy        128 PASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQDYQLGFTL 187
              |  |||||::|   |         |: : ::  :||  :   | :  |          | |:
Db         10 PESPPLSVIVALSLSASPV----TIDSSAAATTVPSFV---FSDGRKLN-------GATV 55

Qy        188 IWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNAT---------LIDSWVD 238
              : : |                   |  |: |   :|  :|| :::           || |||
Db         56 LLRYV-----------------GRSAKKLPDFYG--NNAFDSSQVSILCINMKIDEWVD 95

Qy        239 IA-IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPA 297
              | :|    ||  | |     ::  |  | :| |: |::||| :|| |   |
Db         96 YASVF--SSGSEFENAC--GRVDKYLESSTFLVGHSLSIADVAIWSALAGTG-------- 143

Qy        298 NVQRW 302
               |||
Db        144 --QRW 146
```

RESULT 3
G82441
probable glutathione S-transferase VCA0584 [imported] - Vibrio cholerae (strain N16961 s
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: G82441
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: G82441
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-222 <HEI>
A;Cross-references: UNIPROT:Q9KM05; GB:AE004389; GB:AE003853; NID:g9657989; PIDN:AAF9648
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VCA0584
A;Map position: 2
C;Superfamily: hypothetical protein b2302

Query Match 6.3%; Score 105.5; DB 2; Length 222;
Best Local Similarity 25.7%; Pred. No. 0.17;
Matches 28; Conservative 22; Mismatches 38; Indels 21; Gaps 4;

```
Qy        197 MKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIAIFQLKEGSSKEKAAVF 256
              : | :  :  |: |: |:    :  | :|        || :             ||   :|
Db         97 LMFQMSGVGPMMGQANV---FYRYFPEKIQPA----IDRY-----------QKEGRRLF 137

Qy        257 RSMNSALGKSPWLAGNELTVADVVL--WSVLQQIGGCSVTVPANVQRWM 303
                |:  | ::|:|||:| |:||:    |  : :  | |:    ::||||
Db        138 EVMDGQLAQNPYLAGDEYTIADIATFPWVRIHEWSGISIDGLTHLQRWM 186
```

RESULT 4
A29036
glutathione transferase (EC 2.5.1.18) Yb3 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C;Accession: A29036
R;Abramovitz, M.; Listowsky, I.
J. Biol. Chem. 262, 7770-7773, 1987
A;Title: Selective expression of a unique glutathione S-transferase Yb3 gene in rat brai

A;Reference number: A29036; MUID:87222405; PMID:3584141
A;Accession: A29036
A;Molecule type: mRNA
A;Residues: 1-218 <ABR>
A;Cross-references: UNIPROT:P08009; GB:J02744; NID:g204512; PIDN:AAA41292.1; PID:g20451
C;Superfamily: glutathione transferase
C;Keywords: transferase

Query Match 5.8%; Score 97.5; DB 2; Length 218;
Best Local Similarity 25.6%; Pred. No. 0.78;
Matches 33; Conservative 24; Mismatches 41; Indels 31; Gaps 5;

```
Qy        179 QDYQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNAT------- 231
              : ::||  | : |:|     : |      |     | |:|    |:|||    |
Db         49 EKFKLG--LDFPNLP-----YLIDGSHKITQSNAILRYL----GRKHNLCGETEEERIRV 97

Qy        232 ------LIDSWVDIA-------IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVAD 278
                   |:|: : :|         :|| |  ::   : |  :  ||| || ||:::|  |
Db         98 DILENQLMDNRMVLARLCYNPDFEKLKPGYLEQLPGMMRLYSEFLGKRPWFAGDKITFVD 157

Qy        279 VVLWSVLQQ 287
               : : ||::
Db        158 FIAYDVLER 166
```

RESULT 5
JC5517
Gu/RNA helicase II binding protein - human
C;Species: Homo sapiens (man)
C;Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 05-Nov-1999
C;Accession: JC5517
R;Valdez, B.C.; Henning, D.; Perlaky, L.; Busch, R.K.; Busch, H.
Biochem. Biophys. Res. Commun. 234, 335-340, 1997
A;Title: Cloning and characterization of Gu/RH-II binding protein.
A;Reference number: JC5517; MUID:97320420; PMID:9177271
A;Accession: JC5517
A;Molecule type: mRNA
A;Residues: 1-645 <VAL>
A;Cross-references: GB:U78524; NID:g1696006; PIDN:AAB58488.1; PID:g1696007
A;Note: it is uncertain whether Met-4 or Met-6 is the initiator
C;Comment: This protein is localized to the nucleus and interacts with Gu/RNA helicase I
C;Keywords: phosphoprotein
F;50-58/Region: nuclear location signal
F;362-374/Region: nuclear location signal
F;514-517,551-554,592-595,606-609/Region: 4-residue repeats (N-T-S-L)
F;7,177,322,412,432,460,461,462,470,614,627/Binding site: phosphate (Ser) (covalent) #st
F;65,394,624/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 5.8%; Score 97.5; DB 2; Length 645;
Best Local Similarity 20.5%; Pred. No. 3.7;
Matches 79; Conservative 56; Mismatches 129; Indels 121; Gaps 20;

```
Qy         16 LRV-ELPTCM-YRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQDDILKRLYELKAAV 73
              ||| ||   : |   | ||| :     | |: :      |:: :  :: :| :  |
Db          8 LRVSELQVLLGYAGRNKHGRKHELLTKALHLLKAG--CSPAVQMKIKELYRRRFPQK--- 62

Qy         74 DGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALKDIVINANPASPPL 133
              | || ||| : |:  :  | ||: :              :  :  : :||| ||
Db         63 ------IMTP-ADLSIPNVHSSPMPATLSPS------------TIPQLTYDGHPASSPL 102

Qy        134 ---SLL-VLHRLLCEHF-RVLSTVHTHSSVKSVP-----ENLLK--CFGEQNKKQ----- 176
                 |||   | |   |    |  ||    :: :|     : |:|       | ::
Db        103 LPVSLLGPKHELELPHLTSALHPVHPDIKLQKLPFYDLLDELIKPTSLASDNSQRFRETC 162

Qy        177 ------PRQDYQLGFTLIWKNVPKTQMKFSIQ---------TMCPIEGEG---------- 211
                   |:|  |:  ::   ::  |:  |::|          | || |
Db        163 FAFALTPQQVQQISSSM---DISGTKCDFTVQVQLRFCLSETSCPQEDHFPPNLCVKVNT 219

Qy        212 ------------------------NIARFLFSLFGQKHNAVNATLIDSWV-------DI 239
                                     ||    :    :    |  |:: ||         :
Db        220 KPCSLPGYLPPTKNGVEPKRPSRPINITSLV-----RLSTTVPNTIVVSWTAEIGRNYSM 274

Qy        240 AIFQLKEGSS-------KEKAAVFRSMNSALGKSPWLAGNELTVADVVL-WSVLQQIGGC 291
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:33:04 ; Search time 60.8145 Seconds
(without alignments)
2694.512 Million cell updates/sec

Title: US-10-622-817-3
Perfect score: 1667
Sequence: 1 MPMYQVKPYHGGGAPLRVEL.....RWMRSCENLAPFNTALKLLK 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1667	100.0	320	1	MCA2_HUMAN	Q13155 homo sapien
2	1663	99.8	320	2	Q96CZ5	Q96cz5 homo sapien
3	1464	87.8	320	2	Q8R010	Q8r010 mus musculu
4	1464	87.8	320	2	Q8R2Y6	Q8r2y6 mus musculu
5	1444	86.6	320	1	MCA2_CRIGR	Q9wvm7 cricetulus
6	1260	75.6	280	2	Q8R3V2	Q8r3v2 mus musculu
7	1144.5	68.7	311	2	Q6DK86	Q6dk86 xenopus tro
8	1139.5	68.4	311	2	Q6INU4	Q6inu4 xenopus lae
9	1132.5	67.9	311	2	Q7ZYD7	Q7zyd7 xenopus lae
10	887.5	53.2	321	2	Q7T3C0	Q7t3c0 brachydanio
11	267.5	16.0	340	2	Q7Q7A3	Q7q7a3 anopheles g
12	224	13.4	301	2	Q7KUM5	Q7kum5 drosophila
13	224	13.4	313	2	Q8T060	Q8t060 drosophila
14	224	13.4	322	2	Q6NKM4	Q6nkm4 drosophila
15	214.5	12.9	334	1	MCA2_DROME	Q9vur3 drosophila
16	125.5	7.5	719	2	O82462	O82462 arabidopsis
17	111	6.7	728	2	O65253	O65253 arabidopsis
18	106.5	6.4	880	2	Q8ETT1	Q8ett1 oceanobacil
19	106.5	6.4	913	2	Q6DRB3	Q6drb3 brachydanio
20	106.5	6.4	913	2	Q6PFQ4	Q6pfq4 brachydanio
21	105.5	6.3	222	2	Q9KM05	Q9km05 vibrio chol
22	105	6.3	240	2	Q8BTR1	Q8btr1 mus musculu
23	105	6.3	1512	1	SYEP_MOUSE	Q8cgc7 mus musculu
24	105	6.3	1800	2	Q7PRA2	Q7pra2 anopheles g
25	104.5	6.3	922	2	Q6TEM5	Q6tem5 brachydanio
26	104	6.2	661	2	Q9VB85	Q9vb85 drosophila
27	104	6.2	702	2	Q8IHC6	Q8ihc6 drosophila
28	104	6.2	841	2	Q8IMQ2	Q8imq2 drosophila
29	103	6.2	328	2	Q86X73	Q86x73 homo sapien
30	103	6.2	869	2	Q6PD57	Q6pd57 homo sapien
31	103	6.2	1213	2	Q8FTD2	Q8ftd2 corynebacte
32	102.5	6.1	372	2	Q87XN7	Q87xn7 pseudomonas
33	102.5	6.1	715	2	Q9LIZ8	Q9liz8 oryza sativ
34	102	6.1	559	2	Q8KQ23	Q8kq23 pseudomonas
35	102	6.1	559	2	Q88D25	Q88d25 pseudomonas
36	102	6.1	1486	2	Q6TXE9	Q6txe9 rattus norv
37	101	6.1	1196	2	Q6M580	Q6m580 corynebacte
38	101	6.1	1221	2	Q8NQD1	Q8nqd1 corynebacte
39	100.5	6.0	454	2	Q6UDR4	Q6udr4 mus musculu
40	100.5	6.0	1201	2	Q6NH83	Q6nh83 corynebacte
41	99.5	6.0	651	1	PIA1_MOUSE	O88907 mus musculu
42	99	5.9	203	2	Q6W1I2	Q6w112 rhizobium s
43	98.5	5.9	222	2	Q8EGM4	Q8egm4 shewanella
44	98.5	5.9	582	2	Q66CB1	Q66cb1 yersinia ps
45	97.5	5.8	217	1	GTM3_RAT	P08009 rattus norv

ALIGNMENTS

RESULT 1
MCA2_HUMAN
ID MCA2_HUMAN STANDARD; PRT; 320 AA.
AC Q13155; Q9P1L2;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Multisynthetase complex auxiliary component p38 (JTV-1 protein)
DE (PRO0992).
GN Name=JTV1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96115582; PubMed=8666379;
RA Nicolaides N.C., Kinzler K.W., Vogelstein B.;
RT "Analysis of the 5' region of PMS2 reveals heterogeneous transcripts
RT and a novel overlapping gene.";
RL Genomics 29:329-334(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE OF 197-320 FROM N.A.
RC TISSUE=Fetal liver;
RA Zhang C., Yu Y., Zhang S., Wei H., Zhou G., Ouyang S., Luo L., Bi J.,
RA Liu M., He F.;
RT "Functional prediction of the coding sequences of 121 new genes
RT deduced by analysis of cDNA clones from human fetal liver.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
RN [4]

```
RP   INTERACTION WITH FUBP1.
RX   MEDLINE=22716800; PubMed=12819782; DOI=10.1038/ng1182;
RA   Kim M.J., Park B.-J., Kang Y.-S., Kim H.J., Park J.-H., Kang J.W.,
RA   Lee S.W., Han J.M., Lee H.-W., Kim S.;
RT   "Downregulation of FUSE-binding protein and c-myc by tRNA synthetase
RT   cofactor p38 is required for lung cell differentiation.";
RL   Nat. Genet. 34:330-336(2003).
CC   -!- FUNCTION: Probable core protein of the multisynthetase complex
CC       that serves as a template for the assembly of the supramolecular
CC       structure. Mediates ubiquitination of FUBP1 and its degradation by
CC       the proteasome.
CC   -!- SUBUNIT: Component of the multisynthetase complex which is
CC       comprised of a bifunctional glutamyl-prolyl-tRNA synthetase, the
CC       monospecific isoleucyl, leucyl, glutaminyl, methionyl, lysyl,
CC       arginyl, and aspartyl-tRNA synthetases as well as three auxiliary
CC       proteins, p18, p48 and p43. Binds FUBP1.
CC   -!- SIMILARITY: Contains 1 GST-like domain.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC       frameshift in position 312.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U24169; AAC50391.1; ALT_FRAME.
DR   EMBL; BC002853; AAH02853.1; -.
DR   EMBL; BC010156; AAH10156.1; -.
DR   EMBL; AF116615; AAF71039.1; -.
DR   H-InvDB; HIX0006460; -.
DR   MIM; 600859; -.
DR   InterPro; IPR010987; GST_C_like.
DR   InterPro; IPR004046; GST_Cterm.
DR   Pfam; PF00043; GST_C; 1.
KW   Protein biosynthesis.
SQ   SEQUENCE   320 AA;  35349 MW;  F253726B63C12BAB CRC64;

  Query Match           100.0%;  Score 1667;  DB 1;  Length 320;
  Best Local Similarity  100.0%;  Pred. No. 6e-127;
  Matches  320;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWMRSCENLAPFNTALKLLK 320
              ||||||||||||||||||||
Db        301 RWMRSCENLAPFNTALKLLK 320


RESULT 2
Q96CZ5
ID   Q96CZ5       PRELIMINARY;      PRT;   320 AA.
AC   Q96CZ5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   JTV1.
GN   Name=JTV1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RA   Strausberg R.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC013630; AAH13630.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   320 AA;  35335 MW;  19F14BF758612E08 CRC64;

  Query Match           99.8%;  Score 1663;  DB 2;  Length 320;
  Best Local Similarity  99.7%;  Pred. No. 1.3e-126;
  Matches  319;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIVINANPGSPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWMRSCENLAPFNTALKLLK 320
              ||||||||||||||||||||
Db        301 RWMRSCENLAPFNTALKLLK 320
```

```
RESULT 3
Q8R010
ID   Q8R010      PRELIMINARY;      PRT;   320 AA.
AC   Q8R010;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Jtv1-pending protein.
GN   Name=Jtv1;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N, and FVB/N-3; TISSUE=Mammary tumor;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RA   Strausberg R.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N-3; TISSUE=Mammary tumor;
RA   Strausberg R.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC024410; AAH24410.1; -.
DR   EMBL; BC026972; AAH26972.1; -.
DR   MGD; MGI:2385237; Jtv1.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   320 AA;  35396 MW;  1745D7EE4BC3670D CRC64;

  Query Match             87.8%;  Score 1464;  DB 2;  Length 320;
  Best Local Similarity   87.2%;  Pred. No. 1.8e-110;
  Matches  279;  Conservative   16;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              |||||||||||| |||||||||||||||||| ::  ||  |||||| |  ||||||||||
Db          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              |||||||||||||||||| ||||||||||:|||||||| || ||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
              |||||||||||||||||||||| :|||||||||||:|||||:|||||| :|| | :
Db        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240
              ||||||||||||||||||||:|||||||||||||||||||||||||  ||||||||||

Db        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300
              :|||:|||||||||||||||||||:|||| |||||||||||||||| || |   | |||
Db        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300

Qy        301 RWMRSCENLAPFNTALKLLK 320
              ||::|||||||||:|||:|||
Db        301 RWLKSCENLAPFSTALQLLK 320


RESULT 4
Q8R2Y6
ID   Q8R2Y6      PRELIMINARY;      PRT;   320 AA.
AC   Q8R2Y6;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Jtv1-pending protein.
GN   Name=Jtv1;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CZECH II; TISSUE=Mammary tumor;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CZECH II; TISSUE=Mammary tumor;
RA   Strausberg R.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC026958; AAH26958.1; -.
DR   MGD; MGI:2385237; Jtv1.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   320 AA;  35423 MW;  1C21F1A74C9882B4 CRC64;

  Query Match             87.8%;  Score 1464;  DB 2;  Length 320;
  Best Local Similarity   87.2%;  Pred. No. 1.8e-110;
  Matches  279;  Conservative   16;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              |||||||||||| |||||||||||||||||| ::  ||  |||||| |  ||||||||||
Db          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTNPATDAGHVQETSEPSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              |||||||||||||||||| ||||||||||:|||||||| || ||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
```

```
Db        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Db        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Db        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300

Qy        301 RWMRSCENLAPFNTALKLLK 320

Db        301 RWLKSCENLAPFSTALQLLK 320


RESULT 5
MCA2_CRIGR
ID   MCA2_CRIGR     STANDARD;      PRT;   320 AA.
AC   Q9WVM7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Multisynthetase complex auxiliary component p38.
OS   Cricetulus griseus (Chinese hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Cricetulus.
OX   NCBI_TaxID=10029;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=99096915; PubMed=9878398; DOI=10.1006/jmbi.1998.2316;
RA   Quevillon S., Robinson J.-C., Berthonneau E., Siatecka M., Mirande M.;
RT   "Macromolecular assemblage of aminoacyl-tRNA synthetases:
RT   identification of protein-protein interactions and characterization of
RT   a core protein.";
RL   J. Mol. Biol. 285:183-195(1999).
CC   -!- FUNCTION: Probable core protein of the multisynthetase complex
CC       that serves as a template for the assembly of the supramolecular
CC       structure.
CC   -!- SUBUNIT: Component of the multisynthetase complex which is
CC       comprised of a bifunctional glutamyl-prolyl-tRNA synthetase, the
CC       monospecific isoleucyl, leucyl, glutaminyl, methionyl, lysyl,
CC       arginyl, and aspartyl-tRNA synthetases as well as three auxiliary
CC       proteins, p18, p48 and p43.
CC   -!- SIMILARITY: Contains 1 GST-like domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF072727; AAD38422.1; -.
DR   InterPro; IPR010987; GST_C_like.
DR   InterPro; IPR004046; GST_Cterm.
DR   Pfam; PF00043; GST_C; 1.
KW   Protein biosynthesis.
SQ   SEQUENCE   320 AA;  35433 MW;  6D24E033ABEC810A CRC64;

  Query Match           86.6%;  Score 1444;  DB 1;  Length 320;
  Best Local Similarity  86.6%;  Pred. No. 7.6e-109;
  Matches  277;  Conservative   16;  Mismatches   27;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Db          1 MPMYQVKSYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Db         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPVPLTTNALDLNLVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Db        121 DIVINANPASPPISLLVLHRLLCERYRVLSTVHTHSSVKNVPENLLKCFGEQARKQSRHE 180

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Db        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHSAVNLTLIDSWVDIA 240

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Db        241 MFQLKEGSSKERAAVFRSMNSALGKSPWLVGNELTVADVVLWSVLQQTGDGSGVAPANVQ 300

Qy        301 RWMRSCENLAPFNTALKLLK 320

Db        301 RWLKSCENLVPFSTALQLLK 320


RESULT 6
Q8R3V2
ID   Q8R3V2      PRELIMINARY;      PRT;   280 AA.
AC   Q8R3V2;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   JTV1.
GN   Name=Jtv1;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RA   Strausberg R.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC024480; AAH24480.1; -.
DR   MGD; MGI:2385237; Jtv1.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   280 AA;  31097 MW;  E5BE81498983FD2E CRC64;

  Query Match           75.6%;  Score 1260;  DB 2;  Length 280;
  Best Local Similarity  88.0%;  Pred. No. 5.5e-94;
  Matches  243;  Conservative   14;  Mismatches   19;  Indels    0;  Gaps    0;

Qy         45 QEESNLSLQALESRQDDILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTN 104
```

```
                 || |  |||||||||||||||||||||||||||||||  ||||||||||:||||||||  ||
Db          5 QETSEPSLQALESRQDDILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATN 64

Qy        105 ALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPEN 164
                |||||||||||||||||||||||||||||||||||||||  :|||||||||||||:||||
Db         65 TLDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPEN 124

Qy        165 LLKCFGEQNKKQPRQDYQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQK 224
                |:|||||| :|| | :|||||||||||||||||||:|||||||||||||||||||||||
Db        125 LVKCFGEQARKQSRHEYQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQK 184

Qy        225 HNAVNATLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSV 284
                ||||  ||||||||||:|||:|||||||||||||||||:|||| ||||||||||||||
Db        185 HNAVTLTLIDSWVDIAMFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSV 244

Qy        285 LQQIGGCSVTVPANVQRWMRSCENLAPFNTALKLLK 320
                ||| || |   | |||||::||||||||:|||:|||
Db        245 LQQTGGSSGAAPTNVQRWLKSCENLAPFSTALQLLK 280


RESULT 7
Q6DK86
ID   Q6DK86      PRELIMINARY;      PRT;   311 AA.
AC   Q6DK86;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   MGC69221 protein.
GN   Name=MGC69221;
OS   Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8364;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Gerhard D.S.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC074561; AAH74561.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   311 AA;  34480 MW;  D98F27F73C466154 CRC64;

  Query Match           68.7%;  Score 1144.5;  DB 2;  Length 311;
  Best Local Similarity   68.8%;  Pred. No. 1.5e-84;
  Matches  220;  Conservative   43;  Mismatches   48;  Indels    9;  Gaps    3;

Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||:::|| ||   ::|:|||||||||||         : :::: :: ::|||||||:
Db          1 MPMYRMQPYCGG--EIQVDLPTCMYRLPNVH------QAVSENLEQAADPAIQALESRQE 52

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              |||||||||||||||||||||||||:||| || :       ||:|||||:||||:
Db         53 DILKRLYELKAAVDGLSKMIQTPDADLDVTDIIQTDEQSAPGPVVADLDSVLGKEYGALR 112

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
              ||||||||: |||||||:| |||| ::||| ||||||: |:|| |:||||:| | :||
Db        113 DIVINANPSLPPLSLLILHSLLCERYQVLSAVHTHSSIASIPEPLMKCFGDQLKMKPRHQ 172

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240
              ||||||||||:||| ||||||| |||||||||||||||||| | | ||:|||||| ||| |
Db        173 YQLGFTLIWKDVPKPQMKFSIQNMCPIEGEGNIARFLFSLLGYKFNAINATLIDGWVDTA 232

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300
              ||||:|||||||||| ::||:|||||||| ||||||||:| |  :|| |  | || :||
Db        233 IFQLREGSSKEKAAVLKAMNTALGKSPWLVGNELTVADIVSWCAVQQCGN-STAVPPSVQ 291

Qy        301 RWMRSCENLAPFNTALKLLK 320
              :||:|||||| | : |: ||
Db        292 KWMKSCENLASFKSVLRFLK 311


RESULT 8
Q6INU4
ID   Q6INU4      PRELIMINARY;      PRT;   311 AA.
AC   Q6INU4;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   MGC80304 protein.
GN   Name=MGC80304;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
```

```
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RA   Klein S., Gerhard D.S.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC072178; AAH72178.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   311 AA;  34424 MW;  D5E8325C18D88751 CRC64;

  Query Match             68.4%;  Score 1139.5;  DB 2;  Length 311;
  Best Local Similarity   69.1%;  Pred. No. 3.8e-84;
  Matches  221;  Conservative   40;  Mismatches   50;  Indels    9;  Gaps    3;

Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||:::|| ||   ::|:|||||||||||||  :     | : :: :: ::|||||||:
Db          1 MPMYRMQPYCGG--EIQVDLPTCMYRLPNVHRAT------AENHEQAADPAIQALESRQE 52

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||:|||||| :   :   ||: ||| |||||:
Db         53 DILKRLYELKAAVDGLSKMIQTPDADLDVTDIIQADEQSATGSVVADLDVVLGNDYGALR 112

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
              ||||||||:  |||||:|| |||| ::||| ||||||: |:|| |||||||:| | :|| :
Db        113 DIVINANPSLAPLSLLILHSLLCERYQVLSAVHTHSSITSIPEPLLKCFGDQLKMKPRHE 172

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240
              ||||||||||:||| ||||||| |||||||||| |||||| |   |||||||| ||| |
Db        173 YQLGFTLIWKDVPKPQMKFSIQNMCPIEGEGNIGRFLFSLLGYTFNAVNATLIDGWVDTA 232

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300
              ||||:|||||||||| ::||:|||||||| |||||||:| |  :|| |  |  || |||
Db        233 IFQLREGSSKEKAAVLKAMNTALGKSPWLVGNELTVADIVSWCAVQQCGN-STAVPPNVQ 291

Qy        301 RWMRSCENLAPFNTALKLLK 320
              :||:|||||| | : |: ||
Db        292 KWMKSCENLASFKSVLRFLK 311


RESULT 9
Q7ZYD7
ID   Q7ZYD7      PRELIMINARY;      PRT;   311 AA.
AC   Q7ZYD7;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Jtv1-prov protein.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC043832; AAH43832.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   311 AA;  34647 MW;  2F08C9DA60DD63BE CRC64;

  Query Match             67.9%;  Score 1132.5;  DB 2;  Length 311;
  Best Local Similarity   69.1%;  Pred. No. 1.4e-83;
  Matches  221;  Conservative   39;  Mismatches   51;  Indels    9;  Gaps    4;

Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60
              ||||:::|| ||   ::|:|||||||| ||||       |     | :: :: ::|||||||:
Db          1 MPMYRMQPYCGG--EIQVDLPTCMYRLANVH-----QATSENH-EQAADPAIQALESRQE 52

Qy         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||:||| ||  :   :   ||: ||||:||||:
Db         53 DILKRLYELKAAVDGLSKMIQTPDADLDVTHIIQTDEHSAPGSVVADLDLVLGKEYGALR 112

Qy        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
              ||||||||: |||||:|| |||| ::||| ||||||: |:|| |||||||| | :| | :
Db        113 DIVINANPSLPPLSLLILHSLLCERYQVLSAVHTHSSITSIPEPLLKCFGEQLKMKLRHE 172

Qy        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240
              ||||||||||:||| ||||||| |||||||||| |||||| |   |||||||| ||| |
Db        173 YQLGFTLIWKDVPKPQMKFSIQNMCPIEGEGNIGRFLFSLFGYKFNAVNATLIDGWVDTA 232

Qy        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300
              ||||::|||||||| ::||: ||||||| ||||||||:| |  :|| |  |  || |||
Db        233 IFQLRDGSSKEKAAVLKAMNTTLGKSPWLVGNELTVADIVSWCAVQQCGN-STVVPPNVQ 291

Qy        301 RWMRSCENLAPFNTALKLLK 320
              :||:|||||| | : |: ||
Db        292 KWMKSCENLASFKSVLRFLK 311


RESULT 10
Q7T3C0
ID   Q7T3C0      PRELIMINARY;      PRT;   321 AA.
AC   Q7T3C0;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Zgc:63976.
GN   ORFNames=zgc:63976;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
```

```
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Strausberg R.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC053178; AAH53178.1; -.
DR   ZFIN; ZDB-GENE-040426-2652; zgc:63976.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   321 AA;  34852 MW;  BA8F6B951208244A CRC64;

  Query Match             53.2%;  Score 887.5;  DB 2;  Length 321;
  Best Local Similarity   55.6%;  Pred. No. 1.1e-63;
  Matches  184;  Conservative   50;  Mismatches   76;  Indels   21;  Gaps    8;

Qy          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQE-ESNLSLQALESRQ 59
              ||||||||     | : |:||||||:||||| :  | : |   :|  | : :::||| ||
Db          1 MPMYQVKPV--SPADITVDLPTCMYKLPNVHAQ--GASLGEHALQNGEVDPTVKALEERQ 56

Qy         60 DDILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLT--------TNALDLNSV 111
              |:||::|||||| ||||:| : ||||||| : :        |||            ||: :
Db         57 DEILRKLYELKATVDGLAKTVTTPDADLDASTLAH-----TLTHTPADAVLRGTADLDDL 111

Qy        112 LGKDYGALKDIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGE 171
              |||| |||:|||||||| |||||||| |||: |:|||:|| |||| :||  || | |
Db        112 LGKDLGALRDIVINANPAQPPLSLLVLHALLCQRFQVLSSVHVHSSVSTVPAPLLSCLGP 171

Qy        172 QN-KKQPRQDYQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFG-QKHNAVN 229
              ::     | :|||||||||:| | ||||| | ||||||||:|||: | | :  : |:
Db        172 RHTHSYARHRFQLGFTLIWKDVSKLQMKFSTQNMCPIEGEGNVARFLYRLLGAEPRDPVS 231

Qy        230 ATLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIG 289
              |||:| ||| |:||| || |||:||| |::|:|||:|||| | | ::||:|    : | |
Db        232 ATLMDGWVDTALFQLAEGGSKERAAVLRALNAALGRSPWLLGQEFSLADIVSACCVLQTG 291

Qy        290 GCSVTVPANVQRWMRSCENLAPFNTALKLLK 320
                | : |||||||::||:||  |:    ||:
Db        292 QTS-SAPANVQRWLKSCQNLGYFSCVDPLLQ 321


RESULT 11
Q7Q7A3
ID   Q7Q7A3      PRELIMINARY;      PRT;   340 AA.
AC   Q7Q7A3;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   AgCP5808 (Fragment).
GN   Name=agCG50514; ORFNames=ENSANGG00000011827;
OS   Anopheles gambiae str. PEST.
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX   NCBI_TaxID=180454;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PEST;
RA   Anopheles Genome Sequencing Consortium;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AAAB01008960; EAA10870.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   340 AA;  37467 MW;  FD0D8525CD4EB976 CRC64;

  Query Match             16.0%;  Score 267.5;  DB 2;  Length 340;
  Best Local Similarity   29.2%;  Pred. No. 2.6e-13;
  Matches  100;  Conservative   47;  Mismatches  133;  Indels   63;  Gaps   13;

Qy          3 MYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGP-------APG---AGHV-------- 44
              ||::||       :  ||||||| |  |   |: |       | |   || |
Db         12 MYRLKPVM--NQDVVCELPTCMYTLKPVCAYSHDPKAFSADSATGLAAAGPVDPMSIEID 69

Qy         45 ---QEESNLSLQALESRQDDILKRLYELKAAVDGLSKMIQTPDADLDV-TNIIQADEPTT 100
                |:  |  |: |  ||  :|::| |||       | |     :| :  |   | :|:|
Db         70 NLLQKAENDELRMLAERQQRVLQQLAELK-------KEIMAMRTELKLNANAPPAVQPST 122

Qy        101 -------LTTNALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCEHFRVLSTVH 153
                    |    ::|           |:| |:||:|   | ||| |  |  :   :
Db        123 PLKSKAQLKAEPINLT--------CLQDFVVNASPEYVPYSLLALKNLWKDRLNLQVECF 174

Qy        154 THSSVKSVPE------NLLKCFGEQNKKQPRQDYQLGFTLIWKNV-PKTQMKFSIQTMCP 206
              |||:|  : |      | :  |       ||    : |||||||   |:|  |  :  |
Db        175 THSTVPKLSEEALAFQNAVTASGTAAANLPR----IKVTLIWKNVGAYTEMITSPTSYVP 230

Qy        207 IEGEGNIARFLFSLFGQKHN---AVNATLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSAL 263
              | || || |:|      : |      |   :|| :|      | : : |::  : |:: : |
Db        231 ICGEVNILRYLSRCGPSEFNYEQQDNVDEVDSILDACYLLLNKQNVKQRQQILRTLGAKL 290

Qy        264 GKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQRWMRSC 306
              ||:    | :|:: |:   | ::|:      ||  :|   | |
Db        291 GKAAGFGGADLSLCDIAFTSAVKQV---QRTVAKDVNPNMHKC 330


RESULT 12
Q7KUM5
ID   Q7KUM5      PRELIMINARY;      PRT;   301 AA.
AC   Q7KUM5;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   CG12304-PB.
GN   ORFNames=CG12304;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA   Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
```

```
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA   Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426065; PubMed=12537568;
RA   Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA   Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA   George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA   Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA   Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA   Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT   "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT   melanogaster euchromatic genome sequence.";
RL   Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426070; PubMed=12537573;
RA   Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA   Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA   Ashburner M., Celniker S.E.;
RT   "The transposable elements of the Drosophila melanogaster euchromatin:
RT   a genomics perspective.";
RL   Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfied E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN   [5]
RP   SEQUENCE FROM N.A.
RG   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RG   FlyBase;
RL   Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003530; AAS64998.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   301 AA;  33241 MW;  35224E17FE03F3BE CRC64;

  Query Match           13.4%;  Score 224;  DB 2;  Length 301;
  Best Local Similarity  28.1%;  Pred. No. 7.6e-10;
  Matches   85;  Conservative   47;  Mismatches  120;  Indels   50;  Gaps   10;

Qy        18 VELPTCMYRLPNV--------HGRSYGPAPGAGHVQEESNLSLQALESRQDDILKRLYEL 69
             ::|||||| | ||          | |   :  |      :   |: |   ||: :||:| ||
Db        13 IKLPTCMYPLKNVSLAADSLASGSSTSASTSASTSSCDDTASVAA---RQEKVLKQLEEL 69

Qy        70 KAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALKDIVINANPA 129
             || :  :        | | |    :  : ||   |          |     |:|:||| :|
Db        70 KAQLGQIR-------AGLGVCG--KTFQHTTAFQNG-------GLKEVPLQDVVINGHPN 113

Qy       130 SPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSV-------PENLLKCFGEQNKKQPRQDYQ 182
              | :|| |       : :     |||::  :          || |     |  |:
Db       114 FIPYALLALKNAWRNLYTIDVKTFTHSTMADIGPAAREFEANLAKV--PVNPALPK---- 167

Qy       183 LGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATL---IDSWVDI 239
             :  ||||||   |:|  |      || || || |:|  :  ::   : |   ||  :||
Db       168 ISVTLIWKNCEHTEMISSPTMYVPIYGEVNIIRYLGRVGPAEYRYEGSPLCNEIDLVLDI 227

Qy       240 AIFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANV 299
              |:  : | : |: | ::  | |  :  |::::|||| ::| |       : :||
Db       228 CYQLLRCNTHKTQVAMVRLLDKRLQKQQYFGGSQMSVADVGVYSSL-------IRMPAVT 280

Qy       300 QR 301
             ::
Db       281 EK 282


RESULT 13
Q8T060
ID   Q8T060      PRELIMINARY;      PRT;   313 AA.
AC   Q8T060;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   LD25772p (Fragment).
GN   ORFNames=CG12304;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA   George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA   Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA   Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA   Celniker S.;
RL   Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY069537; AAL39682.2; -.
DR   FlyBase; FBgn0036515; CG12304.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   313 AA;  34522 MW;  5AF9C801877F8531 CRC64;

  Query Match           13.4%;  Score 224;  DB 2;  Length 313;
  Best Local Similarity  28.1%;  Pred. No. 8e-10;
  Matches   85;  Conservative   47;  Mismatches  120;  Indels   50;  Gaps   10;

Qy        18 VELPTCMYRLPNV--------HGRSYGPAPGAGHVQEESNLSLQALESRQDDILKRLYEL 69
             ::|||||| | ||          | |   :  |      :   |: |   ||: :||:| ||
Db        25 IKLPTCMYPLKNVSLAADSLASGSSTSASTSASTSSCDDTASVAA---RQEKVLKQLEEL 81

Qy        70 KAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALKDIVINANPA 129
             || :  :        | | |    :  : ||   |          |     |:|:||| :|
Db        82 KAQLGQIR-------AGLGVCG--KTFQHTTAFQNG-------GLKEVPLQDVVINGHPN 125

Qy       130 SPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSV-------PENLLKCFGEQNKKQPRQDYQ 182
              | :|| |       : :     |||::  :          || |     |  |:
Db       126 FIPYALLALKNAWRNLYTIDVKTFTHSTMADIGPAAREFEANLAKV--PVNPALPK---- 179
```

```
Qy        183 LGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATL---IDSWVDI 239
              :  ||||||   |:|  |      || || || |:|  :   ::    : |   ||  :||
Db        180 ISVTLIWKNCEHTEMISSPTMYVPIYGEVNIIRYLGRVGPAEYRYEGSPLCNEIDLVLDI 239

Qy        240 AIFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLCQIGGCSVTVPANV 299
                |:  : | : |: | ::  | |  :  |::::|||| ::| |       : :||
Db        240 CYQLLRCNTHKTQVAMVRLLDKRLQKQQYFGGSQMSVADVGVYSSL-------IRMPAVT 292

Qy        300 QR 301
              ::
Db        293 EK 294


RESULT 14
Q6NKM4
ID   Q6NKM4      PRELIMINARY;      PRT;   322 AA.
AC   Q6NKM4;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   LP12114p (Fragment).
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Carlson J., Chavez C., Frise E., George R., Pacleb J.,
RA   Park S., Wan K., Yu C., Rubin G.M., Celniker S.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BT012670; AAT08476.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   322 AA;  35528 MW;  82A15927DF928DAE CRC64;

  Query Match             13.4%;  Score 224;  DB 2;  Length 322;
  Best Local Similarity   28.1%;  Pred. No. 8.3e-10;
  Matches   85;  Conservative   47;  Mismatches  120;  Indels   50;  Gaps   10;

Qy         18 VELPTCMYRLPNV--------HGRSYGPAPGAGHVQEESNLSLQALESRQDDILKRLYEL 69
              ::|||||| | ||         | |   :  |     :   |: |   ||: :||:| ||
Db         34 IKLPTCMYPLKNVSLAADSLASGSSTSASTSASTSSCDDTASVAA---RQEKVLKQLEEL 90

Qy         70 KAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALKDIVINANPA 129
              || :  :        | | |    :  : ||   |         |      |:|:||| :|
Db         91 KAQLGQIR-------AGLGVCG--KTFQHTTAFQNG-------GLKEVPLQDVVINGHPN 134

Qy        130 SPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSV-------PENLLKCFGEQNKKQPRQDYQ 182
               | :|| |       : :     |||::  :          || |      |   |:
Db        135 FIPYALLALKNAWRNLYTIDVKTFTHSTMADIGPAAREFEANLAKV--PVNPALPK---- 188

Qy        183 LGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATL---IDSWVDI 239
              :  ||||||   |:|  |      || || || |:|  :   ::    : |   ||  :||
Db        189 ISVTLIWKNCEHTEMISSPTMYVPIYGEVNIIRYLGRVGPAEYRYEGSPLCNEIDLVLDI 248

Qy        240 AIFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANV 299
                |:  : | : |: | ::  | |  :  |::::|||| ::| |       : :||
Db        249 CYQLLRCNTHKTQVAMVRLLDKRLQKQQYFGGSQMSVADVGVYSSL-------IRMPAVT 301

Qy        300 QR 301
              ::
Db        302 EK 303


RESULT 15
MCA2_DROME
ID   MCA2_DROME     STANDARD;      PRT;   334 AA.
AC   Q9VUR3;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   25-JAN-2005 (Rel. 46, Last annotation update)
DE   Probable multisynthetase complex auxiliary component p38.
GN   ORFNames=CG12304;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   GENOME REANNOTATION.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
CC   -!- FUNCTION: Probable core protein of the multisynthetase complex
CC       that serves as a template for the assembly of the supramolecular
CC       structure (By similarity).
CC   -!- SUBUNIT: Component of the multisynthetase complex which is
CC       comprised of a bifunctional glutamyl-prolyl-tRNA synthetase, the
CC       monospecific isoleucyl, leucyl, glutaminyl, methionyl, lysyl,
CC       arginyl, and aspartyl-tRNA synthetases as well as three auxiliary
CC       proteins, p18, p48 and p43 (By similarity).
CC   -!- SIMILARITY: Contains 1 GST-like domain.
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE003530; AAF49612.1; -.
DR   IntAct; Q9VUR3; -.
DR   FlyBase; FBgn0036515; CG12304.
DR   InterPro; IPR010987; GST_C_like.
DR   InterPro; IPR004046; GST_Cterm.
DR   Pfam; PF00043; GST_C; 1.
KW   Protein biosynthesis.
FT   DOMAIN      280    327       GST-like.
SQ   SEQUENCE   334 AA;  36933 MW;  B68FD70AE621990F CRC64;

  Query Match             12.9%;  Score 214.5;  DB 1;  Length 334;
  Best Local Similarity   25.6%;  Pred. No. 5.2e-09;
  Matches   85;  Conservative   50;  Mismatches  120;  Indels   77;  Gaps   10;

Qy         18 VELPTCMYRLPNV----HGRSYGPAPGAGHVQEESNLSLQA------------------ 54
              ::|||||| | ||       : | :  |       |:  |:|
Db         13 IKLPTCMYPLKNVSLAADSLASGSSTSASTSASTSSCKLEANRIDRTGRNAATCALDLDS 72

Qy         55 --------------LESRQDDILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPT 99
                          : :||: :||:| |||| :  :        | | |    :  : |
Db         73 LGRQIQRLLKDDTASVAARQEKVLKQLEELKAQLGQIR-------AGLGVCG--KTFQHT 123

Qy        100 TLTTNALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVK 159
              |    |         |     |:|:||| :|   | :|| |       : :     |||::
Db        124 TAFQNG-------GLKEVPLQDVVINGHPNFIPYALLALKNAWRNLYTIDVKTFTHSTMA 176

Qy        160 SV-------PENLLKCFGEQNKKQPRQDYQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGN 212
              :         || |      |   |:    :  ||||||   |:|  |      || || |
Db        177 DIGPAAREFEANLAKV--PVNPALPK----ISVTLIWKNCEHTEMISSPTMYVPIYGEVN 230

Qy        213 IARFLFSLFGQKHNAVNATL---IDSWVDIAIFQLKEGSSKEKAAVFRSMNSALGKSPWL 269
              | |:|  :   ::    : |   ||  :||    |:  : | : |: | ::  | |  :
Db        231 IIRYLGRVGPAEYRYEGSPLCNEIDLVLDICYQLLRCNTHKTQVAMVRLLDKRLQKQQYF 290

Qy        270 AGNELTVADVVLWSVLQQIGGCSVTVPANVQR 301
               |::::|||| ::| |       : :||  ::
Db        291 GGSQMSVADVGVYSSL-------IRMPAVTEK 315


Search completed: February 23, 2005, 13:57:12
Job time : 61.9811 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:25:39 ; Search time 70.371 Seconds
(without alignments)
1758.725 Million cell updates/sec

Title: US-10-622-817-5
Perfect score: 1655
Sequence: 1 MPMYQVKPYHGGSAPLRVEL.....RWLKSCENLAPFSTALQLLK 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1464	88.5	320	7	ADE58613	Ade58613 Human Pro
2	1464	88.5	320	8	ABM80843	Abm80843 Tumour-as
3	1436	86.8	312	2	AAW25776	Aaw25776 JTV1 prot
4	1436	86.8	312	8	ADR86551	Adr86551 1-312 ami
5	1232	74.4	272	7	ADC10204	Adc10204 Human NOV
6	1069	64.6	229	8	ADR86553	Adr86553 84-312 am
7	735	44.4	161	8	ADR86552	Adr86552 1-161 ami
8	248	15.0	51	5	ABP01502	Abp01502 Human ORF
9	216.5	13.1	334	4	ABB62468	Abb62468 Drosophil
10	148	8.9	263	4	ABG23964	Abg23964 Novel hum
11	126.5	7.6	716	3	AAG41411	Aag41411 Arabidops
12	126.5	7.6	719	3	AAG41410	Aag41410 Arabidops
13	126.5	7.6	748	3	AAG41409	Aag41409 Arabidops
14	106	6.4	1512	4	AAU04349	Aau04349 Mammalian
15	106	6.4	1512	4	AAM78732	Aam78732 Human pro
16	106	6.4	1550	4	AAM79716	Aam79716 Human pro
17	104	6.3	980	8	ADN23683	Adn23683 Bacterial
18	102	6.2	1049	6	ABM67090	Abm67090 Photorhab
19	101.5	6.1	724	4	AAG85046	Aag85046 Shrimp wh
20	99	6.0	691	6	ABU16445	Abu16445 Protein e
21	98	5.9	243	4	AAU58158	Aau58158 Propionib
22	98	5.9	243	6	ABM54677	Abm54677 Propionib
23	98	5.9	644	6	ABJ18952	Abj18952 Pathogen
24	98	5.9	690	4	AAU34345	Aau34345 Staphyloc
25	98	5.9	690	6	ABM71749	Abm71749 Staphyloc
26	98	5.9	985	8	ADN88622	Adn88622 Rat epidi
27	98	5.9	997	8	ADN88620	Adn88620 Rat epidi
28	98	5.9	1004	8	ADN88621	Adn88621 Rat epidi
29	98	5.9	1013	8	ADN88594	Adn88594 Rat epidi
30	98	5.9	1016	8	ADN88619	Adn88619 Rat epidi
31	98	5.9	1032	8	ADN88618	Adn88618 Rat epidi
32	96.5	5.8	1398	8	ABM83058	Abm83058 Human dia
33	96.5	5.8	1398	8	ABM83057	Abm83057 Human dia
34	96.5	5.8	1440	8	ADN03674	Adn03674 Antipsori
35	96.5	5.8	1440	8	ADP54124	Adp54124 Human PRO
36	96.5	5.8	1440	8	ADP23041	Adp23041 PRO polyp
37	96	5.8	2273	2	AAR98811	Aar98811 Erysiphe
38	95	5.7	722	8	ADS20980	Ads20980 Bacterial
39	94	5.7	436	5	ABB54068	Abb54068 Lactococc
40	93.5	5.6	1049	5	ABB91239	Abb91239 Herbicida
41	93.5	5.6	1064	5	AAE25162	Aae25162 RCH1.5 pr
42	93.5	5.6	1090	4	ABG17694	Abg17694 Novel hum
43	93.5	5.6	1090	4	ABG18331	Abg18331 Novel hum
44	93	5.6	204	8	ADS21772	Ads21772 Bacterial
45	93	5.6	559	2	AAR10682	Aar10682 Polyhydro

ALIGNMENTS

RESULT 1
ADE58613
ID ADE58613 standard; protein; 320 AA.
XX
AC ADE58613;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein Q13155, SEQ ID NO 4489.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
DR GENBANK; Q13155.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal

CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 320 AA;

Query Match 88.5%; Score 1464; DB 7; Length 320;
Best Local Similarity 87.2%; Pred. No. 7.5e-142;
Matches 279; Conservative 16; Mismatches 25; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              |||||||||||| |||||||||||||||||| ::  ||  |||||| |  ||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||||| |||||||||||:||||||| || ||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              |||||||||||||||||||||||| :|||||||||||||:|||||:|||||| :|| | :
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              ||||||||||||||||||:|||||||||||||||||||||||||||||  ||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              :|||:|||||||||||||||:|||| |||||||||||||||| || |    | |||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWLKSCENLAPFSTALQLLK 320
              ||::|||||||:|||:|||
Db        301 RWMRSCENLAPFNTALKLLK 320
```

RESULT 2
ABM80843
ID ABM80843 standard; protein; 320 AA.
XX
AC ABM80843;
XX
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) polypeptide PRO81501, SEQ:2179.
XX
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2004030615-A2.
XX
PD 15-APR-2004.
XX
PF 29-SEP-2003; 2003WO-US028547.
XX
PR 02-OCT-2002; 2002US-0414971P.
XX
PA (GETH) GENENTECH INC.
XX
PI Wu TD, Zhang Z, Zhou Y;
XX
DR WPI; 2004-347921/32.
DR N-PSDB; ACN38579.
XX
PT New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
PS Claim 12; SEQ ID NO 2179; 7273pp; English.
XX
CC The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT polypeptide of the invention
XX
SQ Sequence 320 AA;

Query Match 88.5%; Score 1464; DB 8; Length 320;
Best Local Similarity 87.2%; Pred. No. 7.5e-142;
Matches 279; Conservative 16; Mismatches 25; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              |||||||||||| |||||||||||||||||| ::  ||  |||||| |  ||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||||| |||||||||||:||||||| || ||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              |||||||||||||||||||||||| :|||||||||||||:|||||:|||||| :|| | :
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              ||||||||||||||||||:|||||||||||||||||||||||||||||  ||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              :|||:|||||||||||||||:|||| |||||||||||||||| || |    | |||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWLKSCENLAPFSTALQLLK 320
              ||::|||||||:|||:|||
Db        301 RWMRSCENLAPFNTALKLLK 320
```

RESULT 3
AAW25776

ID AAW25776 standard; protein; 312 AA.
XX
AC AAW25776;
XX
DT 19-DEC-1997 (first entry)
XX
DE JTV1 protein.
XX
KW JTV1; hPMS2; probe; detection; chromosome 7; deletion;
KW mismatch repair gene; hereditary non-polyposis colorectal cancer;
KW homologous recombination.
XX
OS Homo sapiens.
XX
PN WO9708312-A1.
XX
PD 06-MAR-1997.
XX
PF 26-AUG-1996; 96WO-US013598.
XX
PR 24-AUG-1995; 95US-00518862.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Vogelstein B, Kinzler KW, Nicolaides NC;
XX
DR WPI; 1997-179269/16.
DR N-PSDB; AAT86182.
XX
PT Novel chromosome 7 gene, JTV1 - used for detecting chromosome 7
PT deletions, and PMS2 promoter activity.
XX
PS Claim 5; Fig 2; 31pp; English.
XX
CC This sequence is JTV1 protein and is encoded by DNA isolated from human
CC chromosome 7. The JTV1 coding sequence is located upstream from hPMS2.
CC JTV1 cDNA can be used as probes to detect chromosome 7 deletions
CC involving JTV1. Due to the overlapping promoter regions, deletions of
CC JTV1 would also affect PMS2 (a mismatch repair gene) expression, leading
CC to hereditary non-polyposis colorectal cancer. JTV1 can also be used to
CC assay activity or competence of the PMS2 promoter region, the presence of
CC JTV1 suggesting that the PMS2 promoter is intact. JTV1 sequences can also
CC be used to guide homologous recombination at the PMS2 locus
XX
SQ Sequence 312 AA;

Query Match 86.8%; Score 1436; DB 2; Length 312;
Best Local Similarity 87.5%; Pred. No. 5.6e-139;
Matches 273; Conservative 14; Mismatches 25; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              ||||||||||||| |||||||||||||||||| ::  ||  |||||| |  ||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||| |||||||||:|||||||| || ||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              |||||||||||||||||||||||| :|||||||||||||:|||||:|||||| :|| | :
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              ||||||||||||||||||:|||||||||||||||||||||||||||  ||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              :|||:|||||||||||||||||:|||| |||||||||||||||| || |   | |||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWLKSCENLAPF 312
              ||::||||||||
Db        301 RWMRSCENLAPF 312
```

RESULT 4
ADR86551
ID ADR86551 standard; protein; 312 AA.
XX
AC ADR86551;
XX
DT 18-NOV-2004 (first entry)
XX
DE 1-312 amino acid sequence of p38/JTV-1 protein.
XX
KW p38/JTV-1; Cytostatic; cancer; leukemia; anticancer.
XX
OS Homo sapiens.
XX
PN EP1454628-A2.
XX
PD 08-SEP-2004.
XX
PF 09-SEP-2003; 2003EP-00020344.
XX
PR 03-MAR-2003; 2003KR-00013058.
XX
PA (UYSE-) UNIV SEOUL NAT IND FOUND.
XX
PI Kim S, Park B;
XX
DR WPI; 2004-627822/61.
DR N-PSDB; ADR86548.
XX
PT New isolated p38/JTV-1 protein, useful as medicament for treating cancer
PT e.g., stomach, liver, blood, bone, pancreatic, skin, head or neck cancer
PT and cutaneous or intraocular melanoma, as well as for screening new
PT anticancer agents.
XX
PS Claim 5; SEQ ID NO 4; 47pp; English.
XX
CC The present invention relates to an isolated p38/JTV-1 protein for use as
CC medicament. The p38/JTV-1 protein or the pharmaceutical composition is
CC useful as medicament for treating breast cancer, large intestinal cancer,
CC lung cancer, small cell lung cancer, stomach cancer, liver cancer, blood
CC cancer, bone cancer, pancreatic cancer, skin cancer, head or neck cancer,
CC cutaneous or intraocular melanoma, uterine sarcoma, ovarian cancer,
CC rectal cancer, anal cancer, colon cancer, fallopian tube carcinoma,
CC endometrial carcinoma, cervical cancer, vulval cancer, vaginal carcinoma,
CC Hodgkin's disease, esophageal cancer, small intestine cancer, endocrine
CC cancer, thyroid cancer, parathyroid cancer, adrenal cancer, soft tissue
CC tumour, urethral cancer, penile cancer, prostate cancer, chronic or acute
CC leukemia, lymphocytic lymphoma, bladder cancer, kidney cancer, ureter
CC cancer, renal cell carcinoma, renal pelvic carcinoma, CNS tumour, primary
CC CNS lymphoma, bone marrow tumour, brain stem nerve gliomas, pituitary
CC adenoma, or their combination. The protein is useful as a target for
CC screening new anticancer agents. The present sequence represents the 1-
CC 312 amino acid sequence of p38/JTV-1 protein.
XX
SQ Sequence 312 AA;

Query Match 86.8%; Score 1436; DB 8; Length 312;
Best Local Similarity 87.5%; Pred. No. 5.6e-139;
Matches 273; Conservative 14; Mismatches 25; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              ||||||||||||| |||||||||||||||||| ::  ||  |||||| |  ||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||| |||||||||:|||||||| || ||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              |||||||||||||||||||||||| :|||||||||||||:|||||:|||||| :|| | :
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
```

```
Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              ||||||||||||||||||||:||||||||||||||||||||||||  ||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              :|||:||||||||||||||||:|||| ||||||||||||||| || |   | |||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWLKSCENLAPF 312
              ||::||||||||
Db        301 RWMRSCENLAPF 312
```

RESULT 5
ADC10204
ID ADC10204 standard; protein; 272 AA.
XX
AC ADC10204;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human NOVX polypeptide SEQ ID NO: 226.
XX
KW cytostatic; antidiabetic; anorectic; cerebroprotective; neuroprotective;
KW antiinflammatory; gene therapy; antisense therapy; thyromimetic; NOVX;
KW pathology; cancer; diabetes; obesity; endocrine disorder; CNS disorder;
KW inflammatory disorder; chromosome mapping; tissue typing;
KW predictive medicine.
XX
OS Homo sapiens.
XX
PN WO2003000842-A2.
XX
PD 03-JAN-2003.
XX
PF 04-JUN-2002; 2002WO-US017443.
XX
PR 04-JUN-2001; 2001US-0295607P.
PR 04-JUN-2001; 2001US-0295661P.
PR 06-JUN-2001; 2001US-0296404P.
PR 06-JUN-2001; 2001US-0296418P.
PR 07-JUN-2001; 2001US-0296575P.
PR 11-JUN-2001; 2001US-0297414P.
PR 12-JUN-2001; 2001US-0295573P.
PR 12-JUN-2001; 2001US-0297567P.
PR 14-JUN-2001; 2001US-0298285P.
PR 15-JUN-2001; 2001US-0298528P.
PR 18-JUN-2001; 2001US-0299133P.
PR 19-JUN-2001; 2001US-0299230P.
PR 21-JUN-2001; 2001US-0299949P.
PR 22-JUN-2001; 2001US-0300177P.
PR 26-JUN-2001; 2001US-0300883P.
PR 28-JUN-2001; 2001US-0301530P.
PR 28-JUN-2001; 2001US-0301550P.
PR 03-JUL-2001; 2001US-0302951P.
PR 31-JUL-2001; 2001US-0308890P.
PR 14-SEP-2001; 2001US-0322297P.
PR 25-SEP-2001; 2001US-0324669P.
PR 03-DEC-2001; 2001US-0337477P.
PR 14-DEC-2001; 2001US-0341562P.
PR 21-FEB-2002; 2002US-0358656P.
PR 21-FEB-2002; 2002US-0359122P.
PR 22-FEB-2002; 2002US-0358978P.
PR 22-FEB-2002; 2002US-0359034P.
PR 22-FEB-2002; 2002US-0359035P.
PR 22-FEB-2002; 2002US-0359121P.
PR 27-FEB-2002; 2002US-0359964P.
PR 01-MAR-2002; 2002US-0360858P.
PR 12-MAR-2002; 2002US-0363430P.
PR 12-MAR-2002; 2002US-0363676P.
PR 10-APR-2002; 2002US-0371346P.
PR 10-MAY-2002; 2002US-0379444P.
PR 04-JUN-2002; 2002US-00379444.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Agee ML, Anderson DW, Berghs C, Casman SJ, Catterton E;
PI Dipippo VA, Edinger SR, Eisen A, Ellerman K, Gangolli EA;
PI Gerlach VL, Gorman L, Guo X, Herrmann JL, Hjalt T, Ji W, Kekuda R;
PI Khramtsov NV, Li L, Liu X, Malyankar UM, Miller CE, Millet I;
PI Ort T, Padigaru M, Patturajan M, Pena CEA, Rastelli L, Rieger DK;
PI Rothenberg ME, Shenoy SG, Shimkets RA, Smithson G, Spaderna SK;
PI Spytek KA, Stone DJ, Vernet CAM, Zhong H, Zhong M, Alsobrook JP;
PI Burgess CE, Lepley DM;
XX
DR WPI; 2003-210149/20.
DR N-PSDB; ADC10203.
XX
PT New isolated NOVX polypeptides and nucleic acid molecules useful for
PT treating, preventing and diagnosing pathological conditions with NOVX-
PT associated disorders, such as cancer, obesity, diabetes and inflammatory
PT or CNS diseases.
XX
PS Claim 1; SEQ ID NO 226; 772pp; English.
XX
CC The invention relates to novel isolated polypeptides, mature form of the
CC polypeptide, a sequence that is 95% identical to the polypeptide or the
CC polypeptide comprising one or more conservative substitutions. The NOVX
CC polypeptide is useful for treating or preventing a pathology associated
CC with the polypeptide e.g. disorders associated with aberrant expression
CC or activity of the polypeptide, such as cancer, diabetes, obesity, and
CC endocrine, CNS and inflammatory disorders. They can also be used in
CC various detection and screening assays, chromosome mapping, tissue typing
CC and predictive medicine. This sequence corresponds to one of the
CC polypeptides of the invention.
XX
SQ Sequence 272 AA;

```
  Query Match             74.4%;  Score 1232;  DB 7;  Length 272;
  Best Local Similarity   88.4%;  Pred. No. 5.1e-118;
  Matches  237;  Conservative   12;  Mismatches   19;  Indels    0;  Gaps    0;

Qy         45 QETSEPSLQALESRQDDILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATN 104
              || |  |||||||||||||||||||||||||||||| ||||||||||:|||||||| ||
Db          5 QEESNLSLQALESRQDDILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTN 64

Qy        105 TLDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPEN 164
              ||||||||||||||||||||||||||||||||||||||| :|||||||||||||:||||
Db         65 ALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPEN 124

Qy        165 LVKCFGEQARKQSRHEYQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQK 224
              |:|||||| :|| | :|||||||||||||||||:||||||||||||||||||||||||||
Db        125 LLKCFGEQNKKQPRQDYQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQK 184

Qy        225 HNAVTLTLIDSWVDIAMFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSV 284
              ||||  ||||||||||:|||:|||||||||||||||||:|||| ||||||||||||||
Db        185 HNAVNATLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSV 244

Qy        285 LQQTGGSSGAAPTNVQRWLKSCENLAPF 312
              ||| || |   | ||||::||||||||
Db        245 LQQIGGCSVTVPANVQRWMRSCENLAPF 272
```

RESULT 6
ADR86553
ID ADR86553 standard; protein; 229 AA.
XX
AC ADR86553;
XX
DT 18-NOV-2004 (first entry)
XX
DE 84-312 amino acid sequence of p38/JTV-1 protein.
XX
KW p38/JTV-1; Cytostatic; cancer; leukemia; anticancer.
XX

OS Homo sapiens.
XX
PN EP1454628-A2.
XX
PD 08-SEP-2004.
XX
PF 09-SEP-2003; 2003EP-00020344.
XX
PR 03-MAR-2003; 2003KR-00013058.
XX
PA (UYSE-) UNIV SEOUL NAT IND FOUND.
XX
PI Kim S, Park B;
XX
DR WPI; 2004-627822/61.
DR N-PSDB; ADR86550.
XX
PT New isolated p38/JTV-1 protein, useful as medicament for treating cancer
PT e.g., stomach, liver, blood, bone, pancreatic, skin, head or neck cancer
PT and cutaneous or intraocular melanoma, as well as for screening new
PT anticancer agents.
XX
PS Claim 5; SEQ ID NO 6; 47pp; English.
XX
CC The present invention relates to an isolated p38/JTV-1 protein for use as
CC medicament. The p38/JTV-1 protein or the pharmaceutical composition is
CC useful as medicament for treating breast cancer, large intestinal cancer,
CC lung cancer, small cell lung cancer, stomach cancer, liver cancer, blood
CC cancer, bone cancer, pancreatic cancer, skin cancer, head or neck cancer,
CC cutaneous or intraocular melanoma, uterine sarcoma, ovarian cancer,
CC rectal cancer, anal cancer, colon cancer, fallopian tube carcinoma,
CC endometrial carcinoma, cervical cancer, vulval cancer, vaginal carcinoma,
CC Hodgkin's disease, esophageal cancer, small intestine cancer, endocrine
CC cancer, thyroid cancer, parathyroid cancer, adrenal cancer, soft tissue
CC tumour, urethral cancer, penile cancer, prostate cancer, chronic or acute
CC leukemia, lymphocytic lymphoma, bladder cancer, kidney cancer, ureter
CC cancer, renal cell carcinoma, renal pelvic carcinoma, CNS tumour, primary
CC CNS lymphoma, bone marrow tumour, brain stem nerve gliomas, pituitary
CC adenoma, or their combination. The protein is useful as a target for
CC screening new anticancer agents. The present sequence represents the 84-
CC 312 amino acid sequence of p38/JTV-1 protein.
XX
SQ Sequence 229 AA;

Query Match 64.6%; Score 1069; DB 8; Length 229;
Best Local Similarity 88.2%; Pred. No. 2.6e-101;
Matches 202; Conservative 12; Mismatches 15; Indels 0; Gaps 0;

```
Qy         84 DADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLC 143
              |||||||||:|||||||| || ||||||||||||||||||||||||||||||||||||||
Db          1 DADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLC 60

Qy        144 ERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHEYQLGFTLIWKNVPKTQMKFSVQT 203
              | :|||||||||||||:|||||:|||||| :|| | :|||||||||||||||||||:||
Db         61 EHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQDYQLGFTLIWKNVPKTQMKFSIQT 120

Qy        204 MCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIAMFQLREGSSKEKAAVFRSMNSAL 263
              |||||||||||||||||||||||||  ||||||||||:|||:||||||||||||||||||
Db        121 MCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSAL 180

Qy        264 GRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQRWLKSCENLAPF 312
              |:|||| |||||||||||||||| || |   | |||||::||||'|||
Db        181 GKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQRWMRSCENLAPF 229
```

RESULT 7
ADR86552
ID ADR86552 standard; protein; 161 AA.
XX
AC ADR86552;
XX
DT 18-NOV-2004 (first entry)
XX
DE 1-161 amino acid sequence of p38/JTV-1 protein.
XX
KW p38/JTV-1; Cytostatic; cancer; leukemia; anticancer.
XX
OS Homo sapiens.
XX
PN EP1454628-A2.
XX
PD 08-SEP-2004.
XX
PF 09-SEP-2003; 2003EP-00020344.
XX
PR 03-MAR-2003; 2003KR-00013058.
XX
PA (UYSE-) UNIV SEOUL NAT IND FOUND.
XX
PI Kim S, Park B;
XX
DR WPI; 2004-627822/61.
DR N-PSDB; ADR86549.
XX
PT New isolated p38/JTV-1 protein, useful as medicament for treating cancer
PT e.g., stomach, liver, blood, bone, pancreatic, skin, head or neck cancer
PT and cutaneous or intraocular melanoma, as well as for screening new
PT anticancer agents.
XX
PS Claim 5; SEQ ID NO 5; 47pp; English.
XX
CC The present invention relates to an isolated p38/JTV-1 protein for use as
CC medicament. The p38/JTV-1 protein or the pharmaceutical composition is
CC useful as medicament for treating breast cancer, large intestinal cancer,
CC lung cancer, small cell lung cancer, stomach cancer, liver cancer, blood
CC cancer, bone cancer, pancreatic cancer, skin cancer, head or neck cancer,
CC cutaneous or intraocular melanoma, uterine sarcoma, ovarian cancer,
CC rectal cancer, anal cancer, colon cancer, fallopian tube carcinoma,
CC endometrial carcinoma, cervical cancer, vulval cancer, vaginal carcinoma,
CC Hodgkin's disease, esophageal cancer, small intestine cancer, endocrine
CC cancer, thyroid cancer, parathyroid cancer, adrenal cancer, soft tissue
CC tumour, urethral cancer, penile cancer, prostate cancer, chronic or acute
CC leukemia, lymphocytic lymphoma, bladder cancer, kidney cancer, ureter
CC cancer, renal cell carcinoma, renal pelvic carcinoma, CNS tumour, primary
CC CNS lymphoma, bone marrow tumour, brain stem nerve gliomas, pituitary
CC adenoma, or their combination. The protein is useful as a target for
CC screening new anticancer agents. The present sequence represents the 1-
CC 161 amino acid sequence of p38/JTV-1 protein.
XX
SQ Sequence 161 AA;

Query Match 44.4%; Score 735; DB 8; Length 161;
Best Local Similarity 88.8%; Pred. No. 4.4e-67;
Matches 143; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              |||||||||||| |||||||||||||||||| ::  ||  |||||| |  |||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||||| ||||||||||:|||||||| || ||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNV 161
              |||||||||||||||||||||||| :|||||||||||||:|
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSV 161
```

RESULT 8
ABP01502
ID ABP01502 standard; protein; 51 AA.
XX
AC ABP01502;
XX
DT 24-JUN-2002 (first entry)
XX

```
DE   Human ORFX protein sequence SEQ ID NO:2986.
XX
KW   Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW   hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW   degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW   cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW   hypertension; hypothyroidism; cholesterol ester storage disease;
KW   immune deficiency; immune disorder; infectious disease;
KW   autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW   myasthenia gravis.
XX
OS   Homo sapiens.
XX
PN   WO200192523-A2.
XX
PD   06-DEC-2001.
XX
PF   29-MAY-2001; 2001WO-US010836.
XX
PR   30-MAY-2000; 2000US-0206132P.
PR   29-AUG-2000; 2000US-0228716P.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach MD;
XX
DR   WPI; 2002-106308/14.
DR   N-PSDB; ABN17254.
XX
PT   Novel human polypeptides and polynucleotides useful for diagnosing,
PT   preventing and treating cardiovascular disease, neurodegenerative,
PT   hyperproliferative disorders and autoimmune disorders.
XX
PS   Disclosure; SEQ ID NO 2986; 1037pp; English.
XX
CC   The present invention describes substantially purified human proteins
CC   (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC   in the specification). ABN15762 to ABN27252 encode the human ORFX
CC   proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC   treating or preventing a pathology associated with an ORFX-associated
CC   disorder in humans, and in the manufacture of a medicament for treating a
CC   syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC   sequences can be used in gene therapy. ORFX sequences can be used in the
CC   treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC   psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC   osteoarthritis, neurodegenerative disorders, disorders related to organ
CC   transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC   lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC   storage disease, various immune deficiencies and disorders, infectious
CC   diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC   arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC   disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC   useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC   bone degenerative disorders, or periodontal disease, and for gut
CC   protection or regeneration and treatment of lung or liver fibrosis,
CC   reperfusion injury in various tissues and conditions resulting from
CC   systemic cytokine damage. N.B. The sequence data for this patent did not
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 51 AA;

  Query Match             15.0%;  Score 248;  DB 5;  Length 51;
  Best Local Similarity   96.1%;  Pred. No. 1.3e-17;
  Matches   49;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         74 DGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALKDIVI 124
              ||||||||||||||||||||||||||||| || ||||||||||||||||||
Db          1 DGLSKMIHTPDADLDVTNILQADEPTTLTTNALDLNSVLGKDYGALKDIVI 51


RESULT 9
ABB62468
ID   ABB62468 standard; protein; 334 AA.
XX
AC   ABB62468;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 14196.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US009231.
XX
PR   23-MAR-2000; 2000US-0191637P.
PR   11-JUL-2000; 2000US-00614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL06571.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signaling and cell-cell
PT   interactions.
XX
PS   Disclosure; SEQ ID NO 14196; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC   ABB72072). The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 334 AA;

  Query Match             13.1%;  Score 216.5;  DB 4;  Length 334;
  Best Local Similarity   26.0%;  Pred. No. 4.8e-13;
  Matches   86;  Conservative   48;  Mismatches  124;  Indels   73;  Gaps    9;

Qy          3 MYQVKPYHGGSAPLRVELPTCMYRLPNV--------HSKTTSPATDAG------------ 42
              ||::|          ::|||||| | ||           :|| :| |
Db          1 MYELKTL---LPQFDIKLPTCMYPLKNVSLAADSLASGSSTSASTSASTSSCKLEANRID 57

Qy         43 -----------------HVQETSEPSLQALESRQDDILKRLYELKAAVDGLSKMIHTPD 84
                               :|   :    :: :||: :||:| |||| :  :
Db         58 RTGRNAATCALDLDSLGRQIQRLLKDDTASVAARQEKVLKQLEELKAQLGQIR------- 110

Qy         85 ADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCE 144
              | | |    :  : ||   |         |     |:|:||| :|   | :|| |
Db        111 AGLGVCG--KTFQHTTAFQNG-------GLKEVPLQDVVINGHPNFIPYALLALKNAWRN 161

Qy        145 RYRVLSTVHTHSSVKNV-------PENLVKCFGEQARKQSRHEYQLGFTLIWKNVPKTQM 197
               | :     |||:: ::         || |      |        ::  ||||||   |:|
Db        162 LYTIDVKTFTHSTMADIGPAAREFEANLAKVPVNPALP------KISVTLIWKNCEHTEM 215

Qy        198 KFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTL---IDSWVDIAMFQLREGSSKEKAA 254
               |     || || || |:|  :   ::      |   ||  :||    ||  : | : |
Db        216 ISSPTMYVPIYGEVNIIRYLGRVGPAEYRYEGSPLCNEIDLVLDICYQLLRCNTHKTQVA 275

Qy        255 VFRSMNSALGRSPWLVGNELTVADVVLWSVL 285
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:36:20 ; Search time 17.81 Seconds
(without alignments)
1341.256 Million cell updates/sec

Title: US-10-622-817-5
Perfect score: 1655
Sequence: 1 MPMYQVKPYHGGSAPLRVEL.....RWLKSCENLAPFSTALQLLK 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1464	88.5	341	4	US-09-949-016-11312	Sequence 11312, A
2	1436	86.8	312	2	US-08-518-862C-2	Sequence 2, Appli
3	106	6.4	1512	3	US-09-443-184-48	Sequence 48, Appl
4	96.5	5.8	1440	3	US-09-357-251-37	Sequence 37, Appl
5	90.5	5.5	359	3	US-09-540-824-2	Sequence 2, Appli
6	90	5.4	559	2	US-08-756-317-10	Sequence 10, Appl
7	90	5.4	559	4	US-09-091-609-4	Sequence 4, Appli
8	90	5.4	2954	4	US-09-150-867-1	Sequence 1, Appli
9	90	5.4	5215	3	US-09-105-537-2	Sequence 2, Appli
10	89	5.4	1056	4	US-09-595-684B-29	Sequence 29, Appl
11	89	5.4	1057	3	US-09-541-782-10	Sequence 10, Appl
12	89	5.4	1057	4	US-09-723-820-10	Sequence 10, Appl
13	89	5.4	1057	4	US-10-270-085-10	Sequence 10, Appl
14	89	5.4	1287	4	US-09-949-016-7826	Sequence 7826, Ap
15	88	5.3	712	4	US-09-489-039A-10736	Sequence 10736, A
16	88	5.3	778	4	US-09-583-110-3930	Sequence 3930, Ap
17	88	5.3	786	4	US-09-107-433-3893	Sequence 3893, Ap
18	86.5	5.2	520	4	US-09-248-796A-17644	Sequence 17644, A
19	86.5	5.2	967	4	US-09-540-236-2449	Sequence 2449, Ap
20	86	5.2	535	4	US-09-489-039A-11461	Sequence 11461, A
21	85.5	5.2	556	4	US-09-134-000C-6329	Sequence 6329, Ap
22	85.5	5.2	608	4	US-09-284-768A-4	Sequence 4, Appli
23	85.5	5.2	657	4	US-09-284-768A-7	Sequence 7, Appli
24	85	5.1	222	4	US-09-248-796A-18146	Sequence 18146, A
25	85	5.1	302	4	US-09-107-532A-6924	Sequence 6924, Ap
26	85	5.1	559	4	US-09-672-749-2	Sequence 2, Appli
27	85	5.1	1971	4	US-09-914-272A-1	Sequence 1, Appli
28	85	5.1	1971	4	US-10-638-333-1	Sequence 1, Appli
29	84	5.1	443	4	US-09-949-016-10598	Sequence 10598, A
30	84	5.1	646	4	US-09-248-796A-18678	Sequence 18678, A
31	84	5.1	1057	4	US-09-428-156B-2	Sequence 2, Appli
32	83.5	5.0	360	4	US-09-949-016-10589	Sequence 10589, A
33	83.5	5.0	475	4	US-09-328-352-5379	Sequence 5379, Ap
34	83.5	5.0	515	4	US-09-489-039A-8208	Sequence 8208, Ap
35	83.5	5.0	519	4	US-09-543-681A-6505	Sequence 6505, Ap
36	83.5	5.0	789	3	US-09-002-285-84	Sequence 84, Appl
37	83.5	5.0	789	4	US-09-589-477-84	Sequence 84, Appl
38	83.5	5.0	789	4	US-10-099-285A-84	Sequence 84, Appl
39	83.5	5.0	959	4	US-09-543-681A-6879	Sequence 6879, Ap
40	83	5.0	434	3	US-09-012-072-4	Sequence 4, Appli
41	83	5.0	434	3	US-09-120-601-4	Sequence 4, Appli
42	83	5.0	448	3	US-09-120-601-6	Sequence 6, Appli
43	83	5.0	559	4	US-09-821-016-1	Sequence 1, Appli
44	83	5.0	559	4	US-10-266-787-1	Sequence 1, Appli
45	83	5.0	767	3	US-08-836-567-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-949-016-11312
; Sequence 11312, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11312
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11312

Query Match 88.5%; Score 1464; DB 4; Length 341;
Best Local Similarity 87.2%; Pred. No. 9e-153;
Matches 279; Conservative 16; Mismatches 25; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              ||||||||||| |||||||||||||||||| ::  ||  |||||| |  ||||||||||
Db         22 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 81

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||||| ||||||||||:|||||||| || ||||||||||||||||
Db         82 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 141

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              |||||||||||||||||||||||||||| :|||||||||||||:|||||:|||||| :|| | :
Db        142 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 201

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              ||||||||||||||||||:|||||||||||||||||||||||||||  ||||||||||
Db        202 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 261

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              :|||:|||||||||||||||:|||| |||||||||||||||| || |   | |||
Db        262 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 321

Qy        301 RWLKSCENLAPFSTALQLLK 320
```

```
                  ||::|||||||||:|||:|||
Db        322 RWMRSCENLAPFNTALKLLK 341


RESULT 2
US-08-518-862C-2
; Sequence 2, Application US/08518862C
; Patent No. 5843757
;  GENERAL INFORMATION:
;    APPLICANT:  Vogelstein, Bert
;    APPLICANT:  Kinzler, Kenneth W.
;    APPLICANT:  Nicolaides, Nicholas C.
;    TITLE OF INVENTION:  Human JTV1 Gene Overlaps PMS2 Gene
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Witcoff, Ltd.
;      STREET:  1001 G Street, N.W.
;      CITY:  Washington, D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/518,862C
;      FILING DATE:  24-AUG-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kagan, Sarah A.
;      REGISTRATION NUMBER:  32,141
;      REFERENCE/DOCKET NUMBER:  01107.49697
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-508-9100
;      TELEFAX:  202-508-9299
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  312 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-518-862C-2

  Query Match             86.8%;  Score 1436;  DB 2;  Length 312;
  Best Local Similarity   87.5%;  Pred. No. 9.6e-150;
  Matches  273;  Conservative   14;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              |||||||||||| |||||||||||||||||| ::  ||  |||||| |  ||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||||| ||||||||||:|||||||| || ||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              |||||||||||||||||||||| :|||||||||||||:|||||:|||||| :|| | :
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              ||||||||||||||||||:|||||||||||||||||||||||||  ||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              :|||:|||||||||||||||:|||| |||||||||||||||| || |   | |||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWLKSCENLAPF 312
              ||::|||||||||
Db        301 RWMRSCENLAPF 312
```

```
RESULT 3
US-09-443-184-48
; Sequence 48, Application US/09443184A
; Patent No. 6372431
; GENERAL INFORMATION:
;  APPLICANT: Cunningham, Mary Jane
;  APPLICANT: Zweiger, Gary
;  APPLICANT: Kaser, Matthew R.
;  APPLICANT: Panzer, Scott
;  APPLICANT: Seilhammer, Jeffrey J.
;  APPLICANT: Yue, Henry
;  APPLICANT: Baughn, Mariah
;  APPLICANT: Azimzai, Yalda
;  APPLICANT: Lal, Preeti
;  TITLE OF INVENTION: MAMMALIAN TOXICOLOGICAL RESPONSE MARKERS
;  FILE REFERENCE: PC-0007 US
;  CURRENT APPLICATION NUMBER: US/09/443,184A
;  CURRENT FILING DATE: 1999-11-19
;  NUMBER OF SEQ ID NOS: 138
;  SOFTWARE: PERL Program
; SEQ ID NO 48
;   LENGTH: 1512
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6372431 2302721CD1
US-09-443-184-48

  Query Match             6.4%;  Score 106;  DB 3;  Length 1512;
  Best Local Similarity   21.8%;  Pred. No. 0.096;
  Matches   44;  Conservative   33;  Mismatches   69;  Indels   56;  Gaps    8;

Qy        124 INANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHEYQL 183
              :  |   |||  |:                  :|::| :: |    |: ::   |
Db          6 LTVNSGDPPLGALL------------------AVEHVKDD-VSISVEEGKENILH---- 41

Qy        184 GFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFL------FSLFGQKHNAVTLTLIDSWV 237
              : :||  | :               :| |:|        |:|   | :  | || |:
Db         42 ----VSENVIFTDV-------------NSILRYLARVATTAGLYGS--NLMEHTEIDHWL 82

Qy        238 DIAMFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQ-----QTGGSS 292
              : :  :|   |      :    :|  |     :|||| |::||: :|: |:     |
Db         83 EFSATKL--SSCDSFTSTINELNHCLSLRTYLVGNSLSLADLCVWATLKGNAAWQEQLKQ 140

Qy        293 GAAPTNVQRWLKSCENLAPFST 314
              || :|:||    |    | :
Db        141 KKAPVHVKRWFGFLEAQQAFQS 162


RESULT 4
US-09-357-251-37
; Sequence 37, Application US/09357251
; Patent No. 6271441
; GENERAL INFORMATION:
;  APPLICANT: Falco, S. Carl
;  APPLICANT: Famodu, Layo O.
;  APPLICANT: Orozco, Buddy
;  APPLICANT: Schwaber, James S.
;  TITLE OF INVENTION: Plant Aminoacyl-tRNA Synthetase
;  FILE REFERENCE: BB-1193
;  CURRENT APPLICATION NUMBER: US/09/357,251
;  CURRENT FILING DATE: 1999-07-20
;  EARLIER APPLICATION NUMBER: 60/093,530
;  EARLIER FILING DATE: July 21, 1998
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 37
;   LENGTH: 1440
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-357-251-37
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:57:26 ; Search time 50.2443 Seconds
(without alignments)
2084.158 Million cell updates/sec

Title: US-10-622-817-5
Perfect score: 1655
Sequence: 1 MPMYQVKPYHGGSAPLRVEL.....RWLKSCENLAPFSTALQLLK 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1380268 seqs, 327241040 residues

Total number of hits satisfying chosen parameters: 1380268

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1436	86.8	312	16	US-10-463-676-4	Sequence 4, Appli
2	1069	64.6	229	16	US-10-463-676-6	Sequence 6, Appli
3	735	44.4	161	16	US-10-463-676-5	Sequence 5, Appli
4	105.5	6.4	925	16	US-10-437-963-126132	Sequence 126132,
5	104	6.3	980	15	US-10-369-493-6336	Sequence 6336, Ap
6	103.5	6.3	468	15	US-10-424-599-260388	Sequence 260388,
7	99.5	6.0	914	16	US-10-437-963-153870	Sequence 153870,
8	99	6.0	691	15	US-10-282-122A-44369	Sequence 44369, A
9	98	5.9	690	9	US-09-815-242-5841	Sequence 5841, Ap
10	96.5	5.8	1440	10	US-09-864-464-37	Sequence 37, Appl
11	96.5	5.8	1440	16	US-10-786-720-31	Sequence 31, Appl
12	96	5.8	569	16	US-10-437-963-123904	Sequence 123904,
13	95	5.7	722	15	US-10-369-493-10013	Sequence 10013, A
14	93.5	5.6	456	15	US-10-425-114-63166	Sequence 63166, A
15	93	5.6	204	15	US-10-369-493-10805	Sequence 10805, A
16	93	5.6	1398	16	US-10-437-963-180124	Sequence 180124,
17	91	5.5	201	15	US-10-369-493-13717	Sequence 13717, A
18	91	5.5	455	15	US-10-425-114-42633	Sequence 42633, A
19	91	5.5	805	15	US-10-424-599-191796	Sequence 191796,
20	91	5.5	1319	16	US-10-437-963-197783	Sequence 197783,
21	91	5.5	1788	16	US-10-437-963-197780	Sequence 197780,
22	90.5	5.5	419	15	US-10-424-599-260387	Sequence 260387,
23	90.5	5.5	578	15	US-10-418-861B-32	Sequence 32, Appl
24	90.5	5.5	661	15	US-10-369-493-2558	Sequence 2558, Ap
25	90.5	5.5	716	15	US-10-369-493-2175	Sequence 2175, Ap
26	90	5.4	559	9	US-09-364-847-21	Sequence 21, Appl
27	90	5.4	856	9	US-09-364-847-33	Sequence 33, Appl
28	90	5.4	856	9	US-09-364-847-35	Sequence 35, Appl
29	90	5.4	1500	16	US-10-437-963-111265	Sequence 111265,
30	90	5.4	5215	9	US-09-861-289-2	Sequence 2, Appli
31	90	5.4	5215	9	US-09-860-846-2	Sequence 2, Appli
32	90	5.4	5215	10	US-09-988-384B-2	Sequence 2, Appli
33	90	5.4	5215	10	US-09-836-821-2	Sequence 2, Appli
34	90	5.4	5215	14	US-10-271-889-45	Sequence 45, Appl
35	89.5	5.4	203	15	US-10-369-493-8604	Sequence 8604, Ap
36	89	5.4	407	15	US-10-424-599-267113	Sequence 267113,
37	89	5.4	1056	15	US-10-282-174-472	Sequence 472, App
38	89	5.4	1056	15	US-10-282-174-474	Sequence 474, App
39	89	5.4	1056	15	US-10-282-174-476	Sequence 476, App
40	89	5.4	1056	17	US-10-600-009-472	Sequence 472, App
41	89	5.4	1056	17	US-10-600-009-474	Sequence 474, App
42	89	5.4	1056	17	US-10-600-009-476	Sequence 476, App
43	89	5.4	1057	16	US-10-714-796-4	Sequence 4, Appli
44	88.5	5.3	656	15	US-10-369-493-18559	Sequence 18559, A
45	88.5	5.3	799	16	US-10-437-963-112603	Sequence 112603,

ALIGNMENTS

RESULT 1
US-10-463-676-4
; Sequence 4, Application US/10463676
; Publication No. US20040175375A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Sunghoon
; APPLICANT: Park, Bum-Joon
; TITLE OF INVENTION: Method for Treating Cancer Using P38/JTV-1 and Method
; TITLE OF INVENTION: for Screening Pharmaceutical Composition for Treating Cancer
; FILE REFERENCE: 012679-091
; CURRENT APPLICATION NUMBER: US/10/463,676
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: KR 10-2003-13058
; PRIOR FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)...(312)
; OTHER INFORMATION: 1-312 amino acid sequence of p38/JTV-1
US-10-463-676-4

Query Match 86.8%; Score 1436; DB 16; Length 312;
Best Local Similarity 87.5%; Pred. No. 2.6e-137;
Matches 273; Conservative 14; Mismatches 25; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              |||||||||||| |||||||||||||||||| ::  ||  |||||| |  ||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||||| ||||||||||:|||||||| || ||||||||||||||
```

```
Db          61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy         121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
               ||||||||||||||||||||||| :|||||||||||||:|||||:|||||| :|| | :
Db         121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy         181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
               ||||||||||||||||||||:|||||||||||||||||||||||||||  ||||||||||
Db         181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy         241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
               :|||:||||||||||||||||:|||| |||||||||||||| || |    | |||
Db         241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy         301 RWLKSCENLAPF 312
               ||::|||||||
Db         301 RWMRSCENLAPF 312


RESULT 2
US-10-463-676-6
; Sequence 6, Application US/10463676
; Publication No. US20040175375A1
; GENERAL INFORMATION:
;  APPLICANT: Kim, Sunghoon
;  APPLICANT:  Park, Bum-Joon
;  TITLE OF INVENTION: Method for Treating Cancer Using P38/JTV-1 and Method
;  TITLE OF INVENTION:  for Screening Pharmaceutical Composition for Treating Cancer
;  FILE REFERENCE: 012679-091
;  CURRENT APPLICATION NUMBER: US/10/463,676
;  CURRENT FILING DATE:  2003-06-18
;  PRIOR APPLICATION NUMBER: KR 10-2003-13058
;  PRIOR FILING DATE: 2003-03-03
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 229
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)...(229)
;   OTHER INFORMATION: 84-312 amino acid sequence p38/JTV-1
US-10-463-676-6

  Query Match             64.6%;  Score 1069;  DB 16;  Length 229;
  Best Local Similarity   88.2%;  Pred. No. 4e-100;
  Matches  202;  Conservative   12;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          84 DADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLC 143
               |||||||||:|||||||| || ||||||||||||||||||||||||||||||||||||||
Db           1 DADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLC 60

Qy         144 ERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHEYQLGFTLIWKNVPKTQMKFSVQT 203
               | :|||||||||||||:|||||:|||||| :|| | :|||||||||||||||||||:||
Db          61 EHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQDYQLGFTLIWKNVPKTQMKFSIQT 120

Qy         204 MCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIAMFQLREGSSKEKAAVFRSMNSAL 263
               ||||||||||||||||||||||||  ||||||||||:|||:|||||||||||||||||||
Db         121 MCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSAL 180

Qy         264 GRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQRWLKSCENLAPF 312
               |:|||| |||||||||||||||| || |    | |||||::||||||||
Db         181 GKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQRWMRSCENLAPF 229


RESULT 3
US-10-463-676-5
; Sequence 5, Application US/10463676
; Publication No. US20040175375A1
; GENERAL INFORMATION:
;  APPLICANT: Kim, Sunghoon
;  APPLICANT:  Park, Bum-Joon
;  TITLE OF INVENTION: Method for Treating Cancer Using P38/JTV-1 and Method
;  TITLE OF INVENTION:  for Screening Pharmaceutical Composition for Treating Cancer
;  FILE REFERENCE: 012679-091
;  CURRENT APPLICATION NUMBER: US/10/463,676
;  CURRENT FILING DATE:  2003-06-18
;  PRIOR APPLICATION NUMBER: KR 10-2003-13058
;  PRIOR FILING DATE: 2003-03-03
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 161
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)...(161)
;   OTHER INFORMATION: 1-161 amino acid sequence of p38/JTV-1
US-10-463-676-5

  Query Match             44.4%;  Score 735;  DB 16;  Length 161;
  Best Local Similarity   88.8%;  Pred. No. 2.5e-66;
  Matches  143;  Conservative    5;  Mismatches   13;  Indels    0;  Gaps    0;

Qy           1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
               |||||||||||| |||||||||||||||||| ::  ||  |||||| |  |||||||||
Db           1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy          61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
               |||||||||||||||||| ||||||||||:|||||||| || ||||||||||||||||
Db          61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy         121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNV 161
               |||||||||||||||||||||||| :|||||||||||||:|
Db         121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSV 161


RESULT 4
US-10-437-963-126132
; Sequence 126132, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 126132
;   LENGTH: 925
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_28709C.1.pep
US-10-437-963-126132

  Query Match              6.4%;  Score 105.5;  DB 16;  Length 925;
  Best Local Similarity   20.3%;  Pred. No. 0.51;
  Matches   64;  Conservative   46;  Mismatches  107;  Indels   99;  Gaps   13;

Qy          25 YRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQDDI--LKRLYELKAAVDGLSKMIHT 82
               | : |: :| | |           | :| ||||  :  ::  :::  |    :  |   :|
Db         152 YMVDNIIAKVTGPV--------TIDPRLQALYKKTTELVGIEKQSEKLVKILSLGDDVHA 203

Qy          83 PDADLDVTNILQAD--EPTTLATNTLDLNSVLGKDYGA------------LKDIVINAN 127
                 |  : : :|:       |||:    | :  |  | ||             |:||:|: :
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:34:50 ; Search time 13.3213 Seconds
(without alignments)
2311.294 Million cell updates/sec

Title: US-10-622-817-5
Perfect score: 1655
Sequence: 1 MPMYQVKPYHGGSAPLRVEL.....RWLKSCENLAPFSTALQLLK 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	126.5	7.6	719	2	T52043	probable glutamate
2	112	6.8	728	2	T01200	probable glutamate
3	104	6.3	980	2	T30089	probable zinc prot
4	99	6.0	691	2	B89797	glycerol ester hyd
5	98.5	6.0	222	2	G82441	probable glutathio
6	98.5	6.0	377	2	F69172	hypothetical prote
7	98	5.9	690	2	A24545	triacylglycerol li
8	96.5	5.8	1440	1	SYHUQT	multifunctional am
9	95.5	5.8	753	2	T24869	hypothetical prote
10	95	5.7	488	1	H64313	corrinoid/iron-sul
11	95	5.7	552	2	D82421	conserved hypothet
12	95	5.7	2655	2	D96595	probable acetyl-Co
13	94.5	5.7	2471	2	T42977	large tegument pro
14	94	5.7	435	2	T01567	hypothetical prote
15	94	5.7	436	2	C86719	GTP-binding protei
16	93.5	5.6	1064	2	B86465	probable Protein k
17	93	5.6	490	2	D84999	ketol-acid reducto
18	92	5.6	754	2	S37403	transcription fact
19	91.5	5.5	436	2	I51237	translation elonga
20	91	5.5	1265	2	T47626	structural mainten
21	90.5	5.5	359	2	T37921	ras-associated pro
22	90.5	5.5	456	2	C84764	hypothetical prote
23	90.5	5.5	642	2	T39490	transcription init
24	90.5	5.5	716	2	T37830	probable glutamate
25	90	5.4	437	2	I51238	translation elonga
26	90	5.4	559	2	A38604	poly(3-hydroxyalka
27	90	5.4	2954	2	T14156	kinesin-related pr
28	89.5	5.4	317	2	T05528	hypothetical prote
29	89.5	5.4	1597	2	S65053	genome polyprotein
30	89.5	5.4	1714	1	S18644	multifunctional am
31	89.5	5.4	4976	2	T14165	peptide synthetase
32	89	5.4	184	2	S74356	glutathione S-tran
33	89	5.4	763	2	AE2443	penicillin-binding
34	89	5.4	1023	2	T31669	neural zinc finger
35	89	5.4	1056	1	G02157	kinesin-like spind
36	89	5.4	1247	2	A33812	interphotoreceptor
37	88.5	5.3	602	2	AB3542	gtp-binding protei
38	88.5	5.3	656	2	H84206	acetyl-CoA synthet
39	88	5.3	821	2	T24728	hypothetical prote
40	88	5.3	2279	2	T42531	acetyl-CoA carboxy
41	88	5.3	2280	2	T38906	acetyl-CoA carboxy
42	87.5	5.3	474	2	B70597	hypothetical prote
43	87.5	5.3	820	2	H82302	ATP-dependent heli
44	87	5.3	570	2	T15763	hypothetical prote
45	86	5.2	906	2	T45158	pre-mRNA splicing

ALIGNMENTS

RESULT 1
T52043
probable glutamate-tRNA ligase (EC 6.1.1.17) [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004
C;Accession: T52043
R;Day, I.S.; Golovkin, M.; Reddy, A.S.
Biochim. Biophys. Acta 1399, 219-24, 1998
A;Title: Cloning of the cDNA for glutamyl-tRNA synthetase from Arabidopsis thaliana.
A;Reference number: Z24836; MUID:9765600; PMID:9765600
A;Accession: T52043
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-719 <DAY>
A;Cross-references: UNIPROT:O82462; EMBL:AF067773; PIDN:AAC36469.1
C;Superfamily: yeast glutamate-tRNA ligase; glutamine-tRNA ligase homology
C;Keywords: ligase

Query Match 7.6%; Score 126.5; DB 2; Length 719;
Best Local Similarity 28.0%; Pred. No. 0.017;
Matches 49; Conservative 22; Mismatches 57; Indels 47; Gaps 8;

```
Qy        128 PASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHEYQLGFTL 187
              | |||||::|    |          |: : ::   ||  :        || :     | |:
Db         10 PESPPLSVIVALSLSASPV----TIDSSAAATTVPSFVF----SDGRKLN------GATV 55

Qy        188 IWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIAMFQLREG 247
              : : |                  |  |: |   :|  :||   : || ||| |      |
Db         56 LLRYV-----------------GRSAKKLPDFYG--NNAFDSSQIDEWVDYASV-FSSG 94

Qy        248 SSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQRW 302
              | | |     ::  |  | :|||: |::||| :|| |  ||          |||
Db         95 SEFENAC--GRVDKYLESSTFLVGHSLSIADVAIWSALAGTG----------QRW 137
```

RESULT 2
T01200
probable glutamate-tRNA ligase (EC 6.1.1.17) F21E10.12 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: T01200
R;Davidson, S.; Rohlfing, T.; David, M.; O'Brian, D.
submitted to the EMBL Data Library, April 1998
A;Description: The sequence of A. thaliana F21E10.
A;Reference number: Z14258
A;Accession: T01200
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-728 <DAV>
A;Cross-references: UNIPROT:O65253; EMBL:AF058914; NID:g3047074; PID:g3047084; GSPDB:GN0
A;Experimental source: cultivar Columbia
C;Genetics:

```
A;Gene: ATSP:F21E10.12
A;Map position: 5
A;Introns: 47/2; 89/3; 141/1; 503/3; 659/3
C;Superfamily: yeast glutamate-tRNA ligase; glutamine-tRNA ligase homology
C;Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis
F;223-499/Domain: glutamine-tRNA ligase homology <EGL>

  Query Match             6.8%;  Score 112;  DB 2;  Length 728;
  Best Local Similarity   26.6%;  Pred. No. 0.29;
  Matches   49;  Conservative   22;  Mismatches   57;  Indels   56;  Gaps    9;

Qy        128 PASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHEYQLGFTL 187
              | |||||::|   |         |: : ::   ||  :         || :       | |:
Db         10 PESPPLSVIVALSLSASPV----TIDSSAAATTVPSFVF----SDGRKLN------GATV 55

Qy        188 IWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLT---------LIDSWVD 238
              : : |                   |  |: |   :|  :||   :          || |||
Db         56 LLRYV-----------------GRSAKKLPDFYG--NNAFDSSQVSILCINMKIDEWVD 95

Qy        239 IAMFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTN 298
              |       ||  | |     ::  |  | :|||: |::||| :|| |  ||
Db         96 YASV-FSSGSEFENAC--GRVDKYLESSTFLVGHSLSIADVAIWSALAGTG--------- 143

Qy        299 VQRW 302
              |||
Db        144 -QRW 146


RESULT 3
T30089
probable zinc proteinase (EC 3.4.24.-) C02G6.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T30089
R;Bentley, D.; Kemp, K.; Scheet, P.
submitted to the EMBL Data Library, April 1996
A;Description: The sequence of C. elegans cosmid C02G6.
A;Reference number: Z20734
A;Accession: T30089
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-980 <BEN>
A;Cross-references: UNIPROT:Q17592; EMBL:U55372; PIDN:AAA98001.1; GSPDB:GN00023; CESP:C0
A;Experimental source: strain Bristol N2; clone C02G6
C;Genetics:
A;Gene: CESP:C02G6.1
A;Map position: 5
A;Introns: 25/3; 215/2; 266/3; 540/3; 585/3; 786/1; 898/2
C;Superfamily: insulin-degrading enzyme (IDE)
C;Keywords: hydrolase; metalloproteinase; zinc
F;70,74/Binding site: zinc (His) #status predicted
F;73/Active site: Glu #status predicted

  Query Match             6.3%;  Score 104;  DB 2;  Length 980;
  Best Local Similarity   20.1%;  Pred. No. 2.1;
  Matches   86;  Conservative   67;  Mismatches  133;  Indels  142;  Gaps   22;

Qy         10 HGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQDDILKRLYE- 68
              | |   | |||    : : :: | : : |   | :  | : |   ||:  |:|:: :
Db        276 HKGPGSLLVELKRLGW-VNSLKSDSNTIAAGFGILNVTMDLSTGGLEN-VDEIIQLMLNY 333

Qy         69 ---LKAAVDGLSKMIHTPDADL-DVT-NILQADEPTTLATNT------LDLNSVLGKDY- 116
                ||:   |  : ||   ||| ||       ::|  :| |       : :  :| |
Db        334 IGMLKSF--GPQQWIHDELADLSDVKFRFKDKEQPMKMAINIAASLQYIPIEHILSSRYL 391

Qy        117 ------GALKDIVINANPASPPLSLLVLHRLLCERYR---------------------- 147
                   :|:::    |::      :| |:: ::::
Db        392 LTKYEPERIKELLSTLTPSN------MLVRVVSQKFKEQEGNTNEPVYGTEMKVTDISPE 445

Qy        148 -------VLSTVH--THSSVKNVPENLVKCFGEQARKQSRHEYQL-----GFTLIW---- 189
                    | | |   |  ||  | :   ||:: |:  ::|:       |:: :|
Db        446 KMKKYENALKTSHHALHLPEKN--EYIATNFGQKPRESVKNEHPKLISDDGWSRVWFKQD 503

Qy        190 --KNVPKTQMKFSVQT------------------MCPIEGE--------GNIARFLFSL 220
                |:|| : ||:: |                  | |  |        |   :|  |
Db        504 DEYNMPKQETKFALTTPIVSQNPRISLISSLWLWCFCDILSEETYNAALAGLGCQFELSP 563

Qy        221 FG-----------QKHNAVTL-----------------TLIDSWVDIAMFQLREGSSKE 251
              ||           ::| ::||                 :|:  :|   |::   | |
Db        564 FGVQKQSTDGREAERHASLTLHVYGYDEKQPLFVKHLTSCMINFKIDRTRFEVLFESLK- 622

Qy        252 KAAVFRSMNSALGRSPWLVG---NELTVADVVLWSVLQQTGGSSGAAPTNVQRWLKSCEN 308
                    |:: :     |:|:    |:| : | | ||  |            ||| : :  |
Db        623 -----RTLTNHAFSQPYLLTQHYNQLLIVDKV-WSKEQLLAVCDSVTLENVQGFAR--EM 674

Qy        309 LAPFSTAL 316
              |  |   |
Db        675 LQAFHMEL 682


RESULT 4
B89797
glycerol ester hydrolase [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: B89797
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of meticillin-resistant Stapylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: B89797
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-691 <KUR>
A;Cross-references: UNIPROT:Q99WQ6; GB:BA000018; PID:g13700235; PIDN:BAB41533.1; GSPDB:G
A;Experimental source: strain N315
C;Genetics:
A;Gene: geh
C;Superfamily: Staphylococcus triacylglycerol lipase

  Query Match             6.0%;  Score 99;  DB 2;  Length 691;
  Best Local Similarity   20.2%;  Pred. No. 3.4;
  Matches   65;  Conservative   53;  Mismatches  110;  Indels   94;  Gaps   15;

Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDA--------GHVQETSEPSL 52
              :|  :| | |  :||       |:        : ||   :| |         : |:| :|:
Db        172 IPAQKVLPNHDKAAPTSTTPPS--------NDKTAPKSTKAQDATTDKHPNQQDTHQPAH 223

Qy         53 QALESRQDDILKRLYELKAAVDGLSKMI---HTPDADLD--------VTNILQADEPTTL 101
              | ::::||| : |  | |  |  ||| |   ::|:   |         : : |||  | |
Db        224 QIIDAKQDDTV-RQSEQKPQVGDLSKHIDGQNSPEKPTDKNTDNKQLIKDALQA--PKTR 280

Qy        102 ATNTLDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNV 161
              :|      :         | |   : ||   |              :| |   |  | :  |
Db        281 STTNAAAD--------AKKVRPLKANQVQP-----------LNKYPV---VFVHGFLGLV 318

Qy        162 PENLVKCFG-----------EQARKQSRHEYQLGFTLIWKNVPK-TQMKFSVQTMCPIEG 209
              :|    :            |: |||  : :|  :       |  :  :: : ::      |
Db        319 GDNAPALYPNYWGGNKFKVIEELRKQGYNVHQASVSAFGSNYDRAVELYYYIKGGRVDYG 378

Qy        210 EGNIARFLFSLFGQKHNAV-------------------TLTLIDSWVDIAMFQLREGSS 249
              : |::    :|: :  :                   |: |:: :       || | :
Db        379 AAHAAKYGHERYGKTYKGIMPNWEPGKKVHLVGHSMGGQTIRLMEEF-------LRNG-N 430

Qy        250 KEKAAVFRSMNSALGRSPWLVG 271
              ||: |  ::    :  ||   |
Db        431 KEEIAYHKAHGGEI--SPLFTG 450


RESULT 5
G82441
probable glutathione S-transferase VCA0584 [imported] - Vibrio cholerae (strain N16961 s
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:33:04 ; Search time 60.8145 Seconds
(without alignments)
2694.512 Million cell updates/sec

Title: US-10-622-817-5
Perfect score: 1655
Sequence: 1 MPMYQVKPYHGGSAPLRVEL.....RWLKSCENLAPFSTALQLLK 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1655	100.0	320	2	Q8R010	Q8r010 mus musculu
2	1652	99.8	320	2	Q8R2Y6	Q8r2y6 mus musculu
3	1573	95.0	320	1	MCA2_CRIGR	Q9wvm7 cricetulus
4	1464	88.5	320	1	MCA2_HUMAN	Q13155 homo sapien
5	1460	88.2	320	2	Q96CZ5	Q96cz5 homo sapien
6	1412	85.3	280	2	Q8R3V2	Q8r3v2 mus musculu
7	1153.5	69.7	311	2	Q6DK86	Q6dk86 xenopus tro
8	1146.5	69.3	311	2	Q6INU4	Q6inu4 xenopus lae
9	1146.5	69.3	311	2	Q7ZYD7	Q7zyd7 xenopus lae
10	926.5	56.0	321	2	Q7T3C0	Q7t3c0 brachydanio
11	247	14.9	340	2	Q7Q7A3	Q7q7a3 anopheles g
12	224	13.5	301	2	Q7KUM5	Q7kum5 drosophila
13	224	13.5	313	2	Q8T060	Q8t060 drosophila
14	224	13.5	322	2	Q6NKM4	Q6nkm4 drosophila
15	216.5	13.1	334	1	MCA2_DROME	Q9vur3 drosophila
16	126.5	7.6	719	2	O82462	O82462 arabidopsis
17	112	6.8	728	2	O65253	O65253 arabidopsis
18	107	6.5	715	2	Q9LIZ8	Q9liz8 oryza sativ
19	106	6.4	328	2	Q86X73	Q86x73 homo sapien
20	106	6.4	559	2	Q8RPZ6	Q8rpz6 pseudomonas
21	106	6.4	869	2	Q6PD57	Q6pd57 homo sapien
22	106	6.4	1721	2	Q7NID9	Q7nid9 gloeobacter
23	104	6.3	980	2	Q17592	Q17592 caenorhabdi
24	102.5	6.2	420	2	Q8FQQ0	Q8fqq0 corynebacte
25	102	6.2	559	2	Q939A8	Q939a8 pseudomonas
26	102	6.2	1048	2	Q7N3E2	Q7n3e2 photorhabdu
27	101.5	6.1	715	2	Q91B77	Q91b77 white spot
28	101.5	6.1	730	2	Q910F5	Q910f5 white spot
29	101.5	6.1	913	2	Q6DRB3	Q6drb3 brachydanio
30	101.5	6.1	913	2	Q6PFQ4	Q6pfq4 brachydanio
31	100.5	6.1	555	2	Q7MEV7	Q7mev7 vibrio vuln
32	100.5	6.1	555	2	Q8D7U3	Q8d7u3 vibrio vuln
33	100.5	6.1	885	2	Q8K2J4	Q8k2j4 mus musculu
34	99.5	6.0	922	2	Q6TEM5	Q6tem5 brachydanio
35	99	6.0	559	2	Q8RQ67	Q8rq67 pseudomonas
36	99	6.0	690	1	LIP_STAAW	Q8nyc2 staphylococ
37	99	6.0	690	2	Q6GCF1	Q6gcf1 staphylococ
38	99	6.0	691	2	Q79SZ7	Q79sz7 staphylococ
39	99	6.0	691	2	Q99WQ6	Q99wq6 staphylococ
40	99	6.0	691	2	Q7A7P2	Q7a7p2 staphylococ
41	98.5	6.0	222	2	Q9KM05	Q9km05 vibrio chol
42	98.5	6.0	377	2	O26161	O26161 methanobact
43	98.5	6.0	1597	2	Q8V9A2	Q8v9a2 crucifer to
44	98	5.9	240	2	Q8BTR1	Q8btr1 mus musculu
45	98	5.9	690	1	LIP_STAAU	P10335 staphylococ

ALIGNMENTS

RESULT 1
Q8R010
ID Q8R010 PRELIMINARY; PRT; 320 AA.
AC Q8R010;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Jtv1-pending protein.
GN Name=Jtv1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N, and FVB/N-3; TISSUE=Mammary tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N-3; TISSUE=Mammary tumor;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC024410; AAH24410.1; -.
DR EMBL; BC026972; AAH26972.1; -.
DR MGD; MGI:2385237; Jtv1.
DR InterPro; IPR004046; GST_Cterm.
DR InterPro; IPR010987; GST_C_like.
DR Pfam; PF00043; GST_C; 1.
SQ SEQUENCE 320 AA; 35396 MW; 1745D7EE4BC3670D CRC64;

```
  Query Match           100.0%;  Score 1655;  DB 2;  Length 320;
  Best Local Similarity  100.0%;  Pred. No. 2.6e-128;
  Matches  320;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300

Qy        301 RWLKSCENLAPFSTALQLLK 320
              ||||||||||||||||||||
Db        301 RWLKSCENLAPFSTALQLLK 320
```

```
RESULT 2
Q8R2Y6
ID   Q8R2Y6      PRELIMINARY;      PRT;   320 AA.
AC   Q8R2Y6;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Jtv1-pending protein.
GN   Name=Jtv1;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CZECH II; TISSUE=Mammary tumor;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CZECH II; TISSUE=Mammary tumor;
RA   Strausberg R.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC026958; AAH26958.1; -.
DR   MGD; MGI:2385237; Jtv1.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   320 AA;  35423 MW;  1C21F1A74C9882B4 CRC64;
```

```
  Query Match           99.8%;  Score 1652;  DB 2;  Length 320;
  Best Local Similarity  99.7%;  Pred. No. 4.6e-128;
  Matches  319;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTNPATDAGHVQETSEPSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300

Qy        301 RWLKSCENLAPFSTALQLLK 320
              ||||||||||||||||||||
Db        301 RWLKSCENLAPFSTALQLLK 320
```

```
RESULT 3
MCA2_CRIGR
ID   MCA2_CRIGR      STANDARD;      PRT;   320 AA.
AC   Q9WVM7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Multisynthetase complex auxiliary component p38.
OS   Cricetulus griseus (Chinese hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Cricetulus.
OX   NCBI_TaxID=10029;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=99096915; PubMed=9878398; DOI=10.1006/jmbi.1998.2316;
RA   Quevillon S., Robinson J.-C., Berthonneau E., Siatecka M., Mirande M.;
RT   "Macromolecular assemblage of aminoacyl-tRNA synthetases:
RT   identification of protein-protein interactions and characterization of
RT   a core protein.";
RL   J. Mol. Biol. 285:183-195(1999).
CC   -!- FUNCTION: Probable core protein of the multisynthetase complex
CC       that serves as a template for the assembly of the supramolecular
CC       structure.
CC   -!- SUBUNIT: Component of the multisynthetase complex which is
CC       comprised of a bifunctional glutamyl-prolyl-tRNA synthetase, the
CC       monospecific isoleucyl, leucyl, glutaminyl, methionyl, lysyl,
CC       arginyl, and aspartyl-tRNA synthetases as well as three auxiliary
CC       proteins, p18, p48 and p43.
CC   -!- SIMILARITY: Contains 1 GST-like domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF072727; AAD38422.1; -.
DR   InterPro; IPR010987; GST_C_like.
DR   InterPro; IPR004046; GST_Cterm.
DR   Pfam; PF00043; GST_C; 1.
KW   Protein biosynthesis.
SQ   SEQUENCE   320 AA;  35433 MW;  6D24E033ABEC810A CRC64;

  Query Match           95.0%;  Score 1573;  DB 1;  Length 320;
  Best Local Similarity   94.4%;  Pred. No. 1.6e-121;
  Matches  302;  Conservative    6;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPMYQVKSYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||||||||||||||||||||  | || |||| ||||||||||
Db         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPVPLTTNALDLNLVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              ||||||||||:|||||||||||||||||||||||||||||||:|||||||||||||||
Db        121 DIVINANPASPPISLLVLHRLLCERYRVLSTVHTHSSVKNVPENLLKCFGEQARKQSRHE 180

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              ||||||||||||||||||||||||||||||||||||||||||:|| ||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHSAVNLTLIDSWVDIA 240

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              ||||:||||||:|||||||||||||:|||||||||||||||||||||  || || |||
Db        241 MFQLKEGSSKERAAVFRSMNSALGKSPWLVGNELTVADVVLWSVLQQTGDGSGVAPANVQ 300

Qy        301 RWLKSCENLAPFSTALQLLK 320
              |||||||||| ||||||||||
Db        301 RWLKSCENLVPFSTALQLLK 320


RESULT 4
MCA2_HUMAN
ID   MCA2_HUMAN     STANDARD;      PRT;   320 AA.
AC   Q13155; Q9P1L2;
DT   01-NOV-1997 (Rel. 35, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Multisynthetase complex auxiliary component p38 (JTV-1 protein)
DE   (PRO0992).
GN   Name=JTV1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96115582; PubMed=8666379;
RA   Nicolaides N.C., Kinzler K.W., Vogelstein B.;
RT   "Analysis of the 5' region of PMS2 reveals heterogeneous transcripts
RT   and a novel overlapping gene.";
RL   Genomics 29:329-334(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lymph;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   SEQUENCE OF 197-320 FROM N.A.
RC   TISSUE=Fetal liver;
RA   Zhang C., Yu Y., Zhang S., Wei H., Zhou G., Ouyang S., Luo L., Bi J.,
RA   Liu M., He F.;
RT   "Functional prediction of the coding sequences of 121 new genes
RT   deduced by analysis of cDNA clones from human fetal liver.";
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   INTERACTION WITH FUBP1.
RX   MEDLINE=22716800; PubMed=12819782; DOI=10.1038/ng1182;
RA   Kim M.J., Park B.-J., Kang Y.-S., Kim H.J., Park J.-H., Kang J.W.,
RA   Lee S.W., Han J.M., Lee H.-W., Kim S.;
RT   "Downregulation of FUSE-binding protein and c-myc by tRNA synthetase
RT   cofactor p38 is required for lung cell differentiation.";
RL   Nat. Genet. 34:330-336(2003).
CC   -!- FUNCTION: Probable core protein of the multisynthetase complex
CC       that serves as a template for the assembly of the supramolecular
CC       structure. Mediates ubiquitination of FUBP1 and its degradation by
CC       the proteasome.
CC   -!- SUBUNIT: Component of the multisynthetase complex which is
CC       comprised of a bifunctional glutamyl-prolyl-tRNA synthetase, the
CC       monospecific isoleucyl, leucyl, glutaminyl, methionyl, lysyl,
CC       arginyl, and aspartyl-tRNA synthetases as well as three auxiliary
CC       proteins, p18, p48 and p43. Binds FUBP1.
CC   -!- SIMILARITY: Contains 1 GST-like domain.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC       frameshift in position 312.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U24169; AAC50391.1; ALT_FRAME.
DR   EMBL; BC002853; AAH02853.1; -.
DR   EMBL; BC010156; AAH10156.1; -.
DR   EMBL; AF116615; AAF71039.1; -.
DR   H-InvDB; HIX0006460; -.
DR   MIM; 600859; -.
DR   InterPro; IPR010987; GST_C_like.
DR   InterPro; IPR004046; GST_Cterm.
DR   Pfam; PF00043; GST_C; 1.
KW   Protein biosynthesis.
SQ   SEQUENCE   320 AA;  35349 MW;  F253726B63C12BAB CRC64;

  Query Match           88.5%;  Score 1464;  DB 1;  Length 320;
  Best Local Similarity   87.2%;  Pred. No. 1.6e-112;
  Matches  279;  Conservative   16;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              |||||||||||| |||||||||||||||||| ::  ||  |||||| |  ||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||| |||||||||:|||||||| || ||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              |||||||||||||||||||||| :|||||||||||||:|||||:|||||| :|| | :
```

```
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              ||||||||||||||||||||:||||||||||||||||||||||||||||  ||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              :|||:|||||||||||||||||:|||| |||||||||||||||| || |    | |||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWLKSCENLAPFSTALQLLK 320
              ||::|||||||:|||:|||
Db        301 RWMRSCENLAPFNTALKLLK 320


RESULT 5
Q96CZ5
ID   Q96CZ5      PRELIMINARY;      PRT;   320 AA.
AC   Q96CZ5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   JTV1.
GN   Name=JTV1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RA   Strausberg R.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC013630; AAH13630.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   320 AA;  35335 MW;  19F14BF758612E08 CRC64;

  Query Match             88.2%;  Score 1460;  DB 2;  Length 320;
  Best Local Similarity   86.9%;  Pred. No. 3.4e-112;
  Matches  278;  Conservative   16;  Mismatches   26;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              |||||||||||| |||||||||||||||| ::  ||  |||||| |  ||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||| |||||||||||:|||||||| || ||||||||||||||||
```

```
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              |||||||| |||||||||||||| :|||||||||||||:|||||:|||||| :|| | :
Db        121 DIVINANPGSPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              ||||||||||||||||||||:||||||||||||||||||||||||||||  ||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              :|||:|||||||||||||||||:|||| |||||||||||||||| || |    | |||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWLKSCENLAPFSTALQLLK 320
              ||::|||||||:|||:|||
Db        301 RWMRSCENLAPFNTALKLLK 320


RESULT 6
Q8R3V2
ID   Q8R3V2      PRELIMINARY;      PRT;   280 AA.
AC   Q8R3V2;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   JTV1.
GN   Name=Jtv1;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RA   Strausberg R.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC024480; AAH24480.1; -.
DR   MGD; MGI:2385237; Jtv1.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   280 AA;  31097 MW;  E5BE81498983FD2E CRC64;

  Query Match             85.3%;  Score 1412;  DB 2;  Length 280;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-108;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         45 QETSEPSLQALESRQDDILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATN 104
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 QETSEPSLQALESRQDDILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATN 64

Qy        105 TLDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPEN 164
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 TLDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPEN 124

Qy        165 LVKCFGEQARKQSRHEYQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQK 224
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 LVKCFGEQARKQSRHEYQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQK 184

Qy        225 HNAVTLTLIDSWVDIAMFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSV 284
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 HNAVTLTLIDSWVDIAMFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSV 244

Qy        285 LQQTGGSSGAAPTNVQRWLKSCENLAPFSTALQLLK 320
                 ||||||||||||||||||||||||||||||||||
Db        245 LQQTGGSSGAAPTNVQRWLKSCENLAPFSTALQLLK 280


RESULT 7
Q6DK86
ID   Q6DK86      PRELIMINARY;      PRT;   311 AA.
AC   Q6DK86;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   MGC69221 protein.
GN   Name=MGC69221;
OS   Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8364;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Gerhard D.S.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC074561; AAH74561.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   311 AA;  34480 MW;  D98F27F73C466154 CRC64;

  Query Match           69.7%;  Score 1153.5;  DB 2;  Length 311;
  Best Local Similarity   67.5%;  Pred. No. 6.8e-87;
  Matches  216;  Conservative   49;  Mismatches   46;  Indels    9;  Gaps    3;

Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              ||||:::|| ||   ::|:|||||||||||||   :       :::: ::|::|||||||:
Db          1 MPMYRMQPYCGGE--IQVDLPTCMYRLPNVHQAVSE------NLEQAADPAIQALESRQE 52

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||| |||||||||:|:| || :        ||:|||||:||||:
Db         53 DILKRLYELKAAVDGLSKMIQTPDADLDVTDIIQTDEQSAPGPVVADLDSVLGKEYGALR 112

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              ||||||||: |||||:|| |||||||||| |||||||: ::|| |:||||:| : : ||:
Db        113 DIVINANPSLPPLSLLILHSLLCERYQVLSAVHTHSSIASIPEPLMKCFGDQLKMKPRHQ 172

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              ||||||||||:||| |||||:| ||||||||||||||||| | | ||:  |||| ||| |
Db        173 YQLGFTLIWKDVPKPQMKFSIQNMCPIEGEGNIARFLFSLLGYKFNAINATLIDGWVDTA 232

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              :|||||||||||||| ::||:|||:|||||||||||||:| |  :|| | |: | | :||
Db        233 IFQLREGSSKEKAAVLKAMNTALGKSPWLVGNELTVADIVSWCAVQQCGNST-AVPPSVQ 291

Qy        301 RWLKSCENLAPFSTALQLLK 320
              :|:||||||| | : |: ||
Db        292 KWMKSCENLASFKSVLRFLK 311


RESULT 8
Q6INU4
ID   Q6INU4      PRELIMINARY;      PRT;   311 AA.
AC   Q6INU4;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   MGC80304 protein.
GN   Name=MGC80304;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
```

```
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RA   Klein S., Gerhard D.S.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC072178; AAH72178.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   311 AA;  34424 MW;  D5E8325C18D88751 CRC64;

  Query Match             69.3%;  Score 1146.5;  DB 2;  Length 311;
  Best Local Similarity   68.1%;  Pred. No. 2.6e-86;
  Matches  218;  Conservative   44;  Mismatches   49;  Indels    9;  Gaps    3;

Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              ||||:::|| ||   ::|:|||||||||||| |      | : :: ::|::|||||||:
Db          1 MPMYRMQPYCGGE--IQVDLPTCMYRLPNVHRAT------AENHEQAADPAIQALESRQE 52

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||| |||||||||:|:|||| :   :   ||: ||| |||||:
Db         53 DILKRLYELKAAVDGLSKMIQTPDADLDVTDIIQADEQSATGSVVADLDVVLGNDYGALR 112

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              ||||||||:  |||||:|| |||||||||| |||||||: ::|| |:||||:| : : |||
Db        113 DIVINANPSLAPLSLLILHSLLCERYQVLSAVHTHSSITSIPEPLLKCFGDQLKMKPRHE 172

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              |||||||||:||| |||||:| |||||||||| |||||| |   |||  |||| ||| |
Db        173 YQLGFTLIWKDVPKPQMKFSIQNMCPIEGEGNIGRFLFSLLGYTFNAVNATLIDGWVDTA 232

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              :|||||||||||||| ::||:|||:|||||||||||||:| |  :|| | |: | | |||
Db        233 IFQLREGSSKEKAAVLKAMNTALGKSPWLVGNELTVADIVSWCAVQQCGNST-AVPPNVQ 291

Qy        301 RWLKSCENLAPFSTALQLLK 320
              :|:||||||| | : |: ||
Db        292 KWMKSCENLASFKSVLRFLK 311


RESULT 9
Q7ZYD7
ID   Q7ZYD7      PRELIMINARY;      PRT;   311 AA.
AC   Q7ZYD7;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Jtv1-prov protein.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC043832; AAH43832.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   311 AA;  34647 MW;  2F08C9DA60DD63BE CRC64;

  Query Match             69.3%;  Score 1146.5;  DB 2;  Length 311;
  Best Local Similarity   68.1%;  Pred. No. 2.6e-86;
  Matches  218;  Conservative   44;  Mismatches   49;  Indels    9;  Gaps    4;

Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              ||||:::|| ||   ::|:|||||||| |||      ||   | :: ::|::|||||||:
Db          1 MPMYRMQPYCGGE--IQVDLPTCMYRLANVHQ-----ATSENH-EQAADPAIQALESRQE 52

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||| |||||||||:|:| || :   :   ||: ||||:||||:
Db         53 DILKRLYELKAAVDGLSKMIQTPDADLDVTHIIQTDEHSAPGSVVADLDLVLGKEYGALR 112

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              ||||||||: |||||:|| |||||||||| |||||||: ::|| |:|||||| : : |||
Db        113 DIVINANPSLPPLSLLILHSLLCERYQVLSAVHTHSSITSIPEPLLKCFGEQLKMKLRHE 172

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              |||||||||:||| |||||:| |||||||||| |||||||| | |||  |||| ||| |
Db        173 YQLGFTLIWKDVPKPQMKFSIQNMCPIEGEGNIGRFLFSLFGYKFNAVNATLIDGWVDTA 232

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              :||||:|||||||| ::||: ||:|||||||||||||:| |  :|| | |:   | |||
Db        233 IFQLRDGSSKEKAAVLKAMNTTLGKSPWLVGNELTVADIVSWCAVQQCGNST-VVPPNVQ 291

Qy        301 RWLKSCENLAPFSTALQLLK 320
              :|:||||||| | : |: ||
Db        292 KWMKSCENLASFKSVLRFLK 311


RESULT 10
Q7T3C0
ID   Q7T3C0      PRELIMINARY;      PRT;   321 AA.
AC   Q7T3C0;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Zgc:63976.
GN   ORFNames=zgc:63976;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
```

```
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Strausberg R.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC053178; AAH53178.1; -.
DR   ZFIN; ZDB-GENE-040426-2652; zgc:63976.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   321 AA;  34852 MW;  BA8F6B951208244A CRC64;

  Query Match             56.0%;  Score 926.5;  DB 2;  Length 321;
  Best Local Similarity   57.3%;  Pred. No. 4.1e-68;
  Matches  188;  Conservative   48;  Mismatches   77;  Indels   15;  Gaps    7;

Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTS---PATDAGHVQETSEPSLQALES 57
              ||||||||     | : |:||||||:|||||::  |    |   | |    :|:::|||
Db          1 MPMYQVKPV--SPADITVDLPTCMYKLPNVHAQGASLGEHALQNGEV----DPTVKALEE 54

Qy         58 RQDDILKRLYELKAAVDGLSKMIHTPDADLDVTNI---LQADEPTTLATNTLDLNSVLGK 114
              |||:||::|||||| ||||:| : ||||||| : :   |        :   | ||: :|||
Db         55 RQDEILRKLYELKATVDGLAKTVTTPDADLDASTLAHTLTHTPADAVLRGTADLDDLLGK 114

Qy        115 DYGALKDIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFG-EQA 173
              | |||:|||||||||| |||||||| |||:|::|||:|| ||||  ||  |: | |
Db        115 DLGALRDIVINANPAQPPLSLLVLHALLCQRFQVLSSVHVHSSVSIVPAPLLSCLGPRHT 174

Qy        174 RKQSRHEYQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFG-QKHNAVTLTL 232
                :|| :|||||||||:| | ||||| | ||||||||:||||: | | :  : |: ||
Db        175 HSYARHRFQLGFTLIWKDVSKLQMKFSTQNMCPIEGEGNVARFLYFLLGAEPRDPVSATL 234

Qy        233 IDSWVDIAMFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSS 292
              :| ||| |:||| || |||:||| |::|:||||||:| | ::||:|    : ||| :|
Db        235 MDGWVDTALFQLAEGGSKERAAVLRALNAALGRSPWLLGQEFSLADIVSACCVLQTGQTS 294

Qy        293 GAAPTNVQRWLKSCENLAPFSTALQLLK 320
               :|| |||||||||:||  ||    ||:
Db        295 -SAPANVQRWLKSCQNLGYFSCVDPLLQ 321


RESULT 11
Q7Q7A3
ID   Q7Q7A3      PRELIMINARY;      PRT;   340 AA.
AC   Q7Q7A3;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   AgCP5808 (Fragment).
GN   Name=agCG50514; ORFNames=ENSANGG00000011827;
OS   Anopheles gambiae str. PEST.
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX   NCBI_TaxID=180454;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PEST;
RA   Anopheles Genome Sequencing Consortium;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AAAB01008960; EAA10870.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   340 AA;  37467 MW;  FD0D8525CD4EB976 CRC64;

  Query Match             14.9%;  Score 247;  DB 2;  Length 340;
  Best Local Similarity   28.7%;  Pred. No. 6.1e-12;
  Matches   94;  Conservative   49;  Mismatches  118;  Indels   66;  Gaps   13;

Qy          3 MYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSP-------ATD---AGHV-------- 44
              ||::||    :  :  ||||||| |  | : :  |        ||    || |
Db         12 MYRLKPVM--NQDVVCELPTCMYTLKPVCAYSHDPKAFSADSATGLAAAGPVDPMSIEID 69

Qy         45 ---QETSEPSLQALESRQDDILKRLYELKAAVDGLSKMIHTPDADLDV-TNILQADEPTT 100
                 |:     |: |  ||  :|::| |||       | |     :| :  |   | :|:|
Db         70 NLLQKAENDELRMLAERQQRVLQQLAELK-------KEIMAMRTELKLNANAPPAVQPST 122

Qy        101 -------LATNTLDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCERYRVLSTVH 153
                    |    ::|           |:| |:||:|   | ||| |  |  :|  :
Db        123 PLKSKAQLKAEPINLT--------CLQDFVVNASPEYVPYSLLALKNLWKDRLNLQVECF 174

Qy        154 THSSVKNVPE------NLVKCFGEQARKQSRHEYQLGFTLIWKNV-PKTQMKFSVQTMCP 206
              |||:|  : |       | |   |  |     |    :  |||||||   |:|  |  :  |
Db        175 THSTVPKLSEEALAFQNAVTASGTAAANLPR----IKVTLIWKNVGAYTEMITSPTSYVP 230

Qy        207 IEGEGNIARFLFSL------FGQKHNAVTLTLIDSWVDIAMFQLREGSSKEKAAVFRSMN 260
              | || || |:|         : |: |   :  :|| :|      | : : |::  : |::
Db        231 ICGEVNILRYLSRCGPSEFNYEQQDN---VDEVDSILDACYLLLNKQNVKQRQQILRTLG 287

Qy        261 SALGRSPWLVGNELTVADVVLWSVLQQ 287
              : ||::    | :|:: |:   | ::|
Db        288 AKLGKAAGFGGADLSLCDIAFTSAVKQ 314


RESULT 12
Q7KUM5
ID   Q7KUM5      PRELIMINARY;      PRT;   301 AA.
AC   Q7KUM5;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   CG12304-PB.
GN   ORFNames=CG12304;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA   Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
```

```
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA   Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426065; PubMed=12537568;
RA   Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA   Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA   George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA   Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA   Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA   Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT   "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT   melanogaster euchromatic genome sequence.";
RL   Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426070; PubMed=12537573;
RA   Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA   Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA   Ashburner M., Celniker S.E.;
RT   "The transposable elements of the Drosophila melanogaster euchromatin:
RT   a genomics perspective.";
RL   Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfied E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN   [5]
RP   SEQUENCE FROM N.A.
RG   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RG   FlyBase;
RL   Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003530; AAS64998.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   301 AA;  33241 MW;  35224E17FE03F3BE CRC64;

  Query Match             13.5%;  Score 224;  DB 2;  Length 301;
  Best Local Similarity   28.6%;  Pred. No. 4.1e-10;
  Matches   86;  Conservative   47;  Mismatches  122;  Indels   46;  Gaps    9;

Qy          3 MYQVKPYHGGSAPLRVELPTCMYRLPNV--------HSKTTSPATDAGHVQETSEPSLQA 54
              ||::|          ::|||||| | ||           :|| :| |     :|     :
Db          1 MYELKTL---LPQFDIKLPTCMYPLKNVSLAADSLASGSSTSASTSA---STSSCDDTAS 54

Qy         55 LESRQDDILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGK 114
              : :||: :||:| |||| :  :          | | |    :  : ||   |        |
Db         55 VAARQEKVLKQLEELKAQLGQIR-------AGLGVCG--KTFQHTTAFQNG-------GL 98

Qy        115 DYGALKDIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNV-------PENLVK 167
              |:|:||| :|   | :|| |       | :     |||:: ::          || |
Db         99 KEVPLQDVVINGHPNFIPYALLALKNAWRNLYTIDVKTFTHSTMADIGPAAREFEANLAK 158

Qy        168 CFGEQARKQSRHEYQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNA 227
              |         ::  ||||||   |:|  |     || || || |:|  :   ::
Db        159 VPVNPALP------KISVTLIWKNCEHTEMISSPTMYVPIYGEVNIIRYLGRVGPAEYRY 212

Qy        228 VTLTL---IDSWVDIAMFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSV 284
              |   || :||    ||  : | : |: | ::  | :  :  |::::|||| ::|
Db        213 EGSPLCNEIDLVLDICYQLLRCNTHKTQVAMVRLLDKRLQKQQYFGGSQMSVADVGVYSS 272

Qy        285 L 285
              |
Db        273 L 273


RESULT 13
Q8T060
ID   Q8T060      PRELIMINARY;      PRT;   313 AA.
AC   Q8T060;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   LD25772p (Fragment).
GN   ORFNames=CG12304;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA   George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA   Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA   Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA   Celniker S.;
RL   Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY069537; AAL39682.2; -.
DR   FlyBase; FBgn0036515; CG12304.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   313 AA;  34522 MW;  5AF9C801877F8531 CRC64;

  Query Match             13.5%;  Score 224;  DB 2;  Length 313;
  Best Local Similarity   28.6%;  Pred. No. 4.3e-10;
  Matches   86;  Conservative   47;  Mismatches  122;  Indels   46;  Gaps    9;

Qy          3 MYQVKPYHGGSAPLRVELPTCMYRLPNV--------HSKTTSPATDAGHVQETSEPSLQA 54
              ||::|          ::|||||| | ||           :|| :| |     :|     :
Db         13 MYELKTL---LPQFDIKLPTCMYPLKNVSLAADSLASGSSTSASTSA---STSSCDDTAS 66

Qy         55 LESRQDDILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGK 114
              : :||: :||:| |||| :  :          | | |    :  : ||   |        |
Db         67 VAARQEKVLKQLEELKAQLGQIR-------AGLGVCG--KTFQHTTAFQNG-------GL 110

Qy        115 DYGALKDIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNV-------PENLVK 167
              |:|:||| :|   | :|| |       | :     |||:: ::          || |
Db        111 KEVPLQDVVINGHPNFIPYALLALKNAWRNLYTIDVKTFTHSTMADIGPAAREFEANLAK 170
```

```
Qy        168 CFGEQARKQSRHEYQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNA 227
                          ::  ||||||   |:|  |     || || || |:|  :   ::
Db        171 VPVNPALP------KISVTLIWKNCEHTEMISSPTMYVPIYGEVNIIRYLGRVGPAEYRY 224

Qy        228 VTLTL---IDSWVDIAMFQLREGSSKEKAAVFRSMNSALGRSPWLVCNELTVADVVLWSV 284
                 |   ||  :||    ||  : | : |: | ::  | :  :  |::::|||| ::|
Db        225 EGSPLCNEIDLVLDICYQLLRCNTHKTQVAMVRLLDKRLQKQQYFGGSQMSVADVGVYSS 284

Qy        285 L 285
                |
Db        285 L 285


RESULT 14
Q6NKM4
ID   Q6NKM4      PRELIMINARY;      PRT;   322 AA.
AC   Q6NKM4;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   LP12114p (Fragment).
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Carlson J., Chavez C., Frise E., George R., Pacleb J.,
RA   Park S., Wan K., Yu C., Rubin G.M., Celniker S.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BT012670; AAT08476.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   322 AA;  35528 MW;  82A15927DF928DAE CRC64;

  Query Match           13.5%;  Score 224;  DB 2;  Length 322;
  Best Local Similarity   28.6%;  Pred. No. 4.5e-10;
  Matches   86;  Conservative   47;  Mismatches  122;  Indels   46;  Gaps    9;

Qy          3 MYQVKPYHGGSAPLRVELPTCMYRLPNV--------HSKTTSPATDAGHVQETSEPSLQA 54
              ||::|          ::||||||  | ||          :|| :| |     :|     :
Db         22 MYELKTL---LPQFDIKLPTCMYPLKNVSLAADSLASGSSTSASTSA---STSSCDDTAS 75

Qy         55 LESRQDDILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGK 114
               : :||: :||:| |||| :  :        | | |    :  : ||   |        |
Db         76 VAARQEKVLKQLEELKAQLGQIR-------AGLGVCG--KTFQHTTAFQNG-------GL 119

Qy        115 DYGALKDIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNV-------PENLVK 167
                |:|:||| :|   | :|| |       | :     |||:: ::          || |
Db        120 KEVPLQDVVINGHPNFIPYALLALKNAWRNLYTIDVKTFTHSTMADIGPAAREFEANLAK 179

Qy        168 CFGEQARKQSRHEYQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNA 227
                      ::  ||||||   |:|  |     || || |  |:|  :   ::
Db        180 VPVNPALP------KISVTLIWKNCEHTEMISSPTMYVPIYGEVNIIRYLGRVGPAEYRY 233

Qy        228 VTLTL---IDSWVDIAMFQLREGSSKEKAAVFRSMNSALGRSPWLVCNELTVADVVLWSV 284
                 |   ||  :||    ||  : | : |: | ::  | :  :  |::::|||| ::|
Db        234 EGSPLCNEIDLVLDICYQLLRCNTHKTQVAMVRLLDKRLQKQQYFGGSQMSVADVGVYSS 293

Qy        285 L 285
                |
Db        294 L 294


RESULT 15
MCA2_DROME
ID   MCA2_DROME     STANDARD;      PRT;   334 AA.
AC   Q9VUR3;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   25-JAN-2005 (Rel. 46, Last annotation update)
DE   Probable multisynthetase complex auxiliary component p38.
GN   ORFNames=CG12304;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   GENOME REANNOTATION.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
CC   -!- FUNCTION: Probable core protein of the multisynthetase complex
CC       that serves as a template for the assembly of the supramolecular
CC       structure (By similarity).
CC   -!- SUBUNIT: Component of the multisynthetase complex which is
CC       comprised of a bifunctional glutamyl-prolyl-tRNA synthetase, the
CC       monospecific isoleucyl, leucyl, glutaminyl, methionyl, lysyl,
CC       arginyl, and aspartyl-tRNA synthetases as well as three auxiliary
CC       proteins, p18, p48 and p43 (By similarity).
CC   -!- SIMILARITY: Contains 1 GST-like domain.
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE003530; AAF49612.1; -.
DR   IntAct; Q9VUR3; -.
DR   FlyBase; FBgn0036515; CG12304.
DR   InterPro; IPR010987; GST_C_like.
DR   InterPro; IPR004046; GST_Cterm.
DR   Pfam; PF00043; GST_C; 1.
KW   Protein biosynthesis.
FT   DOMAIN      280    327       GST-like.
SQ   SEQUENCE   334 AA;  36933 MW;  B68FD70AE621990F CRC64;

  Query Match             13.1%;  Score 216.5;  DB 1;  Length 334;
  Best Local Similarity   26.0%;  Pred. No. 2e-09;
  Matches   86;  Conservative   48;  Mismatches  124;  Indels   73;  Gaps    9;

Qy          3 MYQVKPYHGGSAPLRVELPTCMYRLPNV--------HSKTTSPATDAG------------ 42
              ||::|          ::|||||| | ||           :|| :| |
Db          1 MYELKTL---LPQFDIKLPTCMYPLKNVSLAADSLASGSSTSASTSASTSSCKLEANRID 57

Qy         43 ------------------HVQETSEPSLQALESRQDDILKRLYELKAAVDGLSKMIHTPD 84
                               :|    :    :: :||: :||:| |||| :  :
Db         58 RTGRNAATCALDLDSLGRQIQRLLKDDTASVAARQEKVLKQLEELKAQLGQIR------- 110

Qy         85 ADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCE 144
              | | |    :  : ||   |        |      |:|:||| :|   | :|| |
Db        111 AGLGVCG--KTFQHTTAFQNG-------GLKEVPLQDVVINGHPNFIPYALLALKNAWRN 161

Qy        145 RYRVLSTVHTHSSVKNV-------PENLVKCFGEQARKQSRHEYQLGFTLIWKNVPKTQM 197
               | :     |||:: ::         || |      |        ::  ||||||   |:|
Db        162 LYTIDVKTFTHSTMADIGPAAREFEANLAKVPVNPALP------KISVTLIWKNCEHTEM 215

Qy        198 KFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTL---IDSWVDIAMFQLREGSSKEKAA 254
                |     || || || |:|  :   ::       |   ||  :||    ||  : | : |
Db        216 ISSPTMYVPIYGEVNIIRYLGRVGPAEYRYEGSPLCNEIDLVLDICYQLLRCNTHKTQVA 275

Qy        255 VFRSMNSALGRSPWLVGNELTVADVVLWSVL 285
              : | ::  | :  :  |::::|||| ::| |
Db        276 MVRLLDKRLQKQQYFGGSQMSVADVGVYSSL 306


Search completed: February 23, 2005, 13:57:14
Job time : 62.9811 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:25:39 ; Search time 70.371 Seconds
(without alignments)
1758.725 Million cell updates/sec

Title: US-10-622-817-6
Perfect score: 1655
Sequence: 1 MPMYQVKPYHGGSAPLRVEL.....RWLKSCENLAPFSTALQLLK 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1464	88.5	320	7	ADE58613	Ade58613 Human Pro
2	1464	88.5	320	8	ABM80843	Abm80843 Tumour-as
3	1436	86.8	312	2	AAW25776	Aaw25776 JTV1 prot
4	1436	86.8	312	8	ADR86551	Adr86551 1-312 ami
5	1232	74.4	272	7	ADC10204	Adc10204 Human NOV
6	1069	64.6	229	8	ADR86553	Adr86553 84-312 am
7	735	44.4	161	8	ADR86552	Adr86552 1-161 ami
8	248	15.0	51	5	ABP01502	Abp01502 Human ORF
9	216.5	13.1	334	4	ABB62468	Abb62468 Drosophil
10	148	8.9	263	4	ABG23964	Abg23964 Novel hum
11	126.5	7.6	716	3	AAG41411	Aag41411 Arabidops
12	126.5	7.6	719	3	AAG41410	Aag41410 Arabidops
13	126.5	7.6	748	3	AAG41409	Aag41409 Arabidops
14	106	6.4	1512	4	AAU04349	Aau04349 Mammalian
15	106	6.4	1512	4	AAM78732	Aam78732 Human pro
16	106	6.4	1550	4	AAM79716	Aam79716 Human pro
17	104	6.3	980	8	ADN23683	Adn23683 Bacterial
18	102	6.2	1049	6	ABM67090	Abm67090 Photorhab
19	101.5	6.1	724	4	AAG85046	Aag85046 Shrimp wh
20	99	6.0	691	6	ABU16445	Abu16445 Protein e
21	98	5.9	243	4	AAU58158	Aau58158 Propionib
22	98	5.9	243	6	ABM54677	Abm54677 Propionib
23	98	5.9	644	6	ABJ18952	Abj18952 Pathogen
24	98	5.9	690	4	AAU34345	Aau34345 Staphyloc
25	98	5.9	690	6	ABM71749	Abm71749 Staphyloc
26	98	5.9	985	8	ADN88622	Adn88622 Rat epidi
27	98	5.9	997	8	ADN88620	Adn88620 Rat epidi
28	98	5.9	1004	8	ADN88621	Adn88621 Rat epidi
29	98	5.9	1013	8	ADN88594	Adn88594 Rat epidi
30	98	5.9	1016	8	ADN88619	Adn88619 Rat epidi
31	98	5.9	1032	8	ADN88618	Adn88618 Rat epidi
32	96.5	5.8	1398	8	ABM83058	Abm83058 Human dia
33	96.5	5.8	1398	8	ABM83057	Abm83057 Human dia
34	96.5	5.8	1440	8	ADN03674	Adn03674 Antipsori
35	96.5	5.8	1440	8	ADP54124	Adp54124 Human PRO
36	96.5	5.8	1440	8	ADP23041	Adp23041 PRO polyp
37	96	5.8	2273	2	AAR98811	Aar98811 Erysiphe
38	95	5.7	722	8	ADS20980	Ads20980 Bacterial
39	94	5.7	436	5	ABB54068	Abb54068 Lactococc
40	93.5	5.6	1049	5	ABB91239	Abb91239 Herbicida
41	93.5	5.6	1064	5	AAE25162	Aae25162 RCH1.5 pr
42	93.5	5.6	1090	4	ABG17694	Abg17694 Novel hum
43	93.5	5.6	1090	4	ABG18331	Abg18331 Novel hum
44	93	5.6	204	8	ADS21772	Ads21772 Bacterial
45	93	5.6	559	2	AAR10682	Aar10682 Polyhydro

ALIGNMENTS

RESULT 1
ADE58613
ID ADE58613 standard; protein; 320 AA.
XX
AC ADE58613;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein Q13155, SEQ ID NO 4489.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
DR GENBANK; Q13155.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal

CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 320 AA;

Query Match 88.5%; Score 1464; DB 7; Length 320;
Best Local Similarity 87.2%; Pred. No. 7.5e-142;
Matches 279; Conservative 16; Mismatches 25; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              |||||||||||| |||||||||||||||||| ::  ||  |||||| |  ||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||| ||||||||||:|||||||| || ||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              |||||||||||||||||||||||| :|||||||||||||:|||||:|||||| :|| | :
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              |||||||||||||||||||:|||||||||||||||||||||||||||||  ||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              :|||:|||||||||||||||||:|||| |||||||||||||||| || |   | |||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWLKSCENLAPFSTALQLLK 320
              ||::|||||||:|||:|||
Db        301 RWMRSCENLAPFNTALKLLK 320
```

RESULT 2
ABM80843
ID ABM80843 standard; protein; 320 AA.
XX
AC ABM80843;
XX
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) polypeptide PRO81501, SEQ:2179.
XX
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2004030615-A2.
XX
PD 15-APR-2004.
XX
PF 29-SEP-2003; 2003WO-US028547.
XX
PR 02-OCT-2002; 2002US-0414971P.
XX
PA (GETH) GENENTECH INC.
XX
PI Wu TD, Zhang Z, Zhou Y;
XX
DR WPI; 2004-347921/32.
DR N-PSDB; ACN38579.
XX
PT New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
PS Claim 12; SEQ ID NO 2179; 7273pp; English.
XX
CC The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT polypeptide of the invention
XX
SQ Sequence 320 AA;

Query Match 88.5%; Score 1464; DB 8; Length 320;
Best Local Similarity 87.2%; Pred. No. 7.5e-142;
Matches 279; Conservative 16; Mismatches 25; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              |||||||||||| |||||||||||||||||| ::  ||  |||||| |  ||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||| ||||||||||:|||||||| || ||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              |||||||||||||||||||||||| :|||||||||||||:|||||:|||||| :|| | :
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              |||||||||||||||||||:|||||||||||||||||||||||||||||  ||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              :|||:|||||||||||||||||:|||| |||||||||||||||| || |   | |||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWLKSCENLAPFSTALQLLK 320
              ||::|||||||:|||:|||
Db        301 RWMRSCENLAPFNTALKLLK 320
```

RESULT 3
AAW25776

ID AAW25776 standard; protein; 312 AA.
XX
AC AAW25776;
XX
DT 19-DEC-1997 (first entry)
XX
DE JTV1 protein.
XX
KW JTV1; hPMS2; probe; detection; chromosome 7; deletion;
KW mismatch repair gene; hereditary non-polyposis colorectal cancer;
KW homologous recombination.
XX
OS Homo sapiens.
XX
PN WO9708312-A1.
XX
PD 06-MAR-1997.
XX
PF 26-AUG-1996; 96WO-US013598.
XX
PR 24-AUG-1995; 95US-00518862.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Vogelstein B, Kinzler KW, Nicolaides NC;
XX
DR WPI; 1997-179269/16.
DR N-PSDB; AAT86182.
XX
PT Novel chromosome 7 gene, JTV1 - used for detecting chromosome 7
PT deletions, and PMS2 promoter activity.
XX
PS Claim 5; Fig 2; 31pp; English.
XX
CC This sequence is JTV1 protein and is encoded by DNA isolated from human
CC chromosome 7. The JTV1 coding sequence is located upstream from hPMS2.
CC JTV1 cDNA can be used as probes to detect chromosome 7 deletions
CC involving JTV1. Due to the overlapping promoter regions, deletions of
CC JTV1 would also affect PMS2 (a mismatch repair gene) expression, leading
CC to hereditary non-polyposis colorectal cancer. JTV1 can also be used to
CC assay activity or competence of the PMS2 promoter region, the presence of
CC JTV1 suggesting that the PMS2 promoter is intact. JTV1 sequences can also
CC be used to guide homologous recombination at the PMS2 locus
XX
SQ Sequence 312 AA;

Query Match 86.8%; Score 1436; DB 2; Length 312;
Best Local Similarity 87.5%; Pred. No. 5.6e-139;
Matches 273; Conservative 14; Mismatches 25; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              |||||||||||| |||||||||||||||||| ::  ||  |||||| |  ||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||| ||||||||||:||||||| || ||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              |||||||||||||||||||||| :||||||||||||:|||||:|||||| :|| | :
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              ||||||||||||||||||:|||||||||||||||||||||||||  ||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              :|||:||||||||||||||||:|||| |||||||||||||||| || |  | |||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWLKSCENLAPF 312
              ||::||||||||
Db        301 RWMRSCENLAPF 312
```

RESULT 4
ADR86551
ID ADR86551 standard; protein; 312 AA.
XX
AC ADR86551;
XX
DT 18-NOV-2004 (first entry)
XX
DE 1-312 amino acid sequence of p38/JTV-1 protein.
XX
KW p38/JTV-1; Cytostatic; cancer; leukemia; anticancer.
XX
OS Homo sapiens.
XX
PN EP1454628-A2.
XX
PD 08-SEP-2004.
XX
PF 09-SEP-2003; 2003EP-00020344.
XX
PR 03-MAR-2003; 2003KR-00013058.
XX
PA (UYSE-) UNIV SEOUL NAT IND FOUND.
XX
PI Kim S, Park B;
XX
DR WPI; 2004-627822/61.
DR N-PSDB; ADR86548.
XX
PT New isolated p38/JTV-1 protein, useful as medicament for treating cancer
PT e.g., stomach, liver, blood, bone, pancreatic, skin, head or neck cancer
PT and cutaneous or intraocular melanoma, as well as for screening new
PT anticancer agents.
XX
PS Claim 5; SEQ ID NO 4; 47pp; English.
XX
CC The present invention relates to an isolated p38/JTV-1 protein for use as
CC medicament. The p38/JTV-1 protein or the pharmaceutical composition is
CC useful as medicament for treating breast cancer, large intestinal cancer,
CC lung cancer, small cell lung cancer, stomach cancer, liver cancer, blood
CC cancer, bone cancer, pancreatic cancer, skin cancer, head or neck cancer,
CC cutaneous or intraocular melanoma, uterine sarcoma, ovarian cancer,
CC rectal cancer, anal cancer, colon cancer, fallopian tube carcinoma,
CC endometrial carcinoma, cervical cancer, vulval cancer, vaginal carcinoma,
CC Hodgkin's disease, esophageal cancer, small intestine cancer, endocrine
CC cancer, thyroid cancer, parathyroid cancer, adrenal cancer, soft tissue
CC tumour, urethral cancer, penile cancer, prostate cancer, chronic or acute
CC leukemia, lymphocytic lymphoma, bladder cancer, kidney cancer, ureter
CC cancer, renal cell carcinoma, renal pelvic carcinoma, CNS tumour, primary
CC CNS lymphoma, bone marrow tumour, brain stem nerve gliomas, pituitary
CC adenoma, or their combination. The protein is useful as a target for
CC screening new anticancer agents. The present sequence represents the 1-
CC 312 amino acid sequence of p38/JTV-1 protein.
XX
SQ Sequence 312 AA;

Query Match 86.8%; Score 1436; DB 8; Length 312;
Best Local Similarity 87.5%; Pred. No. 5.6e-139;
Matches 273; Conservative 14; Mismatches 25; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              |||||||||||| |||||||||||||||||| ::  ||  |||||| |  ||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||| ||||||||||:||||||| || ||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              |||||||||||||||||||||| :||||||||||||:|||||:|||||| :|| | :
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180
```

```
Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              ||||||||||||||||||||:||||||||||||||||||||||||||  ||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              :|||:||||||||||||||||:|||| ||||||||||||||| || |   | |||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWLKSCENLAPF 312
              ||::|||||||
Db        301 RWMRSCENLAPF 312


RESULT 5
ADC10204
ID   ADC10204 standard; protein; 272 AA.
XX
AC   ADC10204;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Human NOVX polypeptide SEQ ID NO: 226.
XX
KW   cytostatic; antidiabetic; anorectic; cerebroprotective; neuroprotective;
KW   antiinflammatory; gene therapy; antisense therapy; thyromimetic; NOVX;
KW   pathology; cancer; diabetes; obesity; endocrine disorder; CNS disorder;
KW   inflammatory disorder; chromosome mapping; tissue typing;
KW   predictive medicine.
XX
OS   Homo sapiens.
XX
PN   WO2003000842-A2.
XX
PD   03-JAN-2003.
XX
PF   04-JUN-2002; 2002WO-US017443.
XX
PR   04-JUN-2001; 2001US-0295607P.
PR   04-JUN-2001; 2001US-0295661P.
PR   06-JUN-2001; 2001US-0296404P.
PR   06-JUN-2001; 2001US-0296418P.
PR   07-JUN-2001; 2001US-0296575P.
PR   11-JUN-2001; 2001US-0297414P.
PR   12-JUN-2001; 2001US-0295573P.
PR   12-JUN-2001; 2001US-0297567P.
PR   14-JUN-2001; 2001US-0298285P.
PR   15-JUN-2001; 2001US-0298528P.
PR   18-JUN-2001; 2001US-0299133P.
PR   19-JUN-2001; 2001US-0299230P.
PR   21-JUN-2001; 2001US-0299949P.
PR   22-JUN-2001; 2001US-0300177P.
PR   26-JUN-2001; 2001US-0300883P.
PR   28-JUN-2001; 2001US-0301530P.
PR   28-JUN-2001; 2001US-0301550P.
PR   03-JUL-2001; 2001US-0302951P.
PR   31-JUL-2001; 2001US-0308890P.
PR   14-SEP-2001; 2001US-0322297P.
PR   25-SEP-2001; 2001US-0324669P.
PR   03-DEC-2001; 2001US-0337477P.
PR   14-DEC-2001; 2001US-0341562P.
PR   21-FEB-2002; 2002US-0358656P.
PR   21-FEB-2002; 2002US-0359122P.
PR   22-FEB-2002; 2002US-0358978P.
PR   22-FEB-2002; 2002US-0359034P.
PR   22-FEB-2002; 2002US-0359035P.
PR   22-FEB-2002; 2002US-0359121P.
PR   27-FEB-2002; 2002US-0359964P.
PR   01-MAR-2002; 2002US-0360858P.
PR   12-MAR-2002; 2002US-0363430P.
PR   12-MAR-2002; 2002US-0363676P.
PR   10-APR-2002; 2002US-0371346P.
PR   10-MAY-2002; 2002US-0379444P.
PR   04-JUN-2002; 2002US-00379444.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Agee ML,  Anderson DW,  Berghs C,  Casman SJ,  Catterton E;
PI   Dipippo VA,  Edinger SR,  Eisen A,  Ellerman K,  Gangolli EA;
PI   Gerlach VL,  Gorman L,  Guo X,  Herrmann JL,  Hjalt T,  Ji W,  Kekuda R;
PI   Khramtsov NV,  Li L,  Liu X,  Malyankar UM,  Miller CE,  Millet I;
PI   Ort T,  Padigaru M,  Patturajan M,  Pena CEA,  Rastelli L,  Rieger DK;
PI   Rothenberg ME,  Shenoy SG,  Shimkets RA,  Smithson G,  Spaderna SK;
PI   Spytek KA,  Stone DJ,  Vernet CAM,  Zhong H,  Zhong M,  Alsobrook JP;
PI   Burgess CE,  Lepley DM;
XX
DR   WPI; 2003-210149/20.
DR   N-PSDB; ADC10203.
XX
PT   New isolated NOVX polypeptides and nucleic acid molecules useful for
PT   treating, preventing and diagnosing pathological conditions with NOVX-
PT   associated disorders, such as cancer, obesity, diabetes and inflammatory
PT   or CNS diseases.
XX
PS   Claim 1; SEQ ID NO 226; 772pp; English.
XX
CC   The invention relates to novel isolated polypeptides, mature form of the
CC   polypeptide, a sequence that is 95% identical to the polypeptide or the
CC   polypeptide comprising one or more conservative substitutions. The NOVX
CC   polypeptide is useful for treating or preventing a pathology associated
CC   with the polypeptide e.g. disorders associated with aberrant expression
CC   or activity of the polypeptide, such as cancer, diabetes, obesity, and
CC   endocrine, CNS and inflammatory disorders. They can also be used in
CC   various detection and screening assays, chromosome mapping, tissue typing
CC   and predictive medicine. This sequence corresponds to one of the
CC   polypeptides of the invention.
XX
SQ   Sequence 272 AA;

  Query Match            74.4%;  Score 1232;  DB 7;  Length 272;
  Best Local Similarity  88.4%;  Pred. No. 5.1e-118;
  Matches  237;  Conservative   12;  Mismatches   19;  Indels    0;  Gaps    0;

Qy         45 QETSEPSLQALESRQDDILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATN 104
              || |  ||||||||||||||||||||||||||||||| ||||||||||:|||||||| ||
Db          5 QEESNLSLQALESRQDDILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTN 64

Qy        105 TLDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPEN 164
              ||||||||||||||||||||||||||||||||||||||| :|||||||||||||:||||
Db         65 ALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPEN 124

Qy        165 LVKCFGEQARKQSRHEYQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQK 224
              |:|||||| :|| | :|||||||||||||||||||:|||||||||||||||||||||||
Db        125 LLKCFGEQNKKQPRQDYQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQK 184

Qy        225 HNAVTLTLIDSWVDIAMFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSV 284
              ||||  ||||||||||:|||:|||||||||||||||||:|||| ||||||||||||||
Db        185 HNAVNATLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSV 244

Qy        285 LQQTGGSSGAAPTNVQRWLKSCENLAPF 312
              ||| || |   | |||||::||||||||
Db        245 LQQIGGCSVTVPANVQRWMRSCENLAPF 272


RESULT 6
ADR86553
ID   ADR86553 standard; protein; 229 AA.
XX
AC   ADR86553;
XX
DT   18-NOV-2004  (first entry)
XX
DE   84-312 amino acid sequence of p38/JTV-1 protein.
XX
KW   p38/JTV-1; Cytostatic; cancer; leukemia; anticancer.
XX
```

OS Homo sapiens.
XX
PN EP1454628-A2.
XX
PD 08-SEP-2004.
XX
PF 09-SEP-2003; 2003EP-00020344.
XX
PR 03-MAR-2003; 2003KR-00013058.
XX
PA (UYSE-) UNIV SEOUL NAT IND FOUND.
XX
PI Kim S, Park B;
XX
DR WPI; 2004-627822/61.
DR N-PSDB; ADR86550.
XX
PT New isolated p38/JTV-1 protein, useful as medicament for treating cancer
PT e.g., stomach, liver, blood, bone, pancreatic, skin, head or neck cancer
PT and cutaneous or intraocular melanoma, as well as for screening new
PT anticancer agents.
XX
PS Claim 5; SEQ ID NO 6; 47pp; English.
XX
CC The present invention relates to an isolated p38/JTV-1 protein for use as
CC medicament. The p38/JTV-1 protein or the pharmaceutical composition is
CC useful as medicament for treating breast cancer, large intestinal cancer,
CC lung cancer, small cell lung cancer, stomach cancer, liver cancer, blood
CC cancer, bone cancer, pancreatic cancer, skin cancer, head or neck cancer,
CC cutaneous or intraocular melanoma, uterine sarcoma, ovarian cancer,
CC rectal cancer, anal cancer, colon cancer, fallopian tube carcinoma,
CC endometrial carcinoma, cervical cancer, vulval cancer, vaginal carcinoma,
CC Hodgkin's disease, esophageal cancer, small intestine cancer, endocrine
CC cancer, thyroid cancer, parathyroid cancer, adrenal cancer, soft tissue
CC tumour, urethral cancer, penile cancer, prostate cancer, chronic or acute
CC leukemia, lymphocytic lymphoma, bladder cancer, kidney cancer, ureter
CC cancer, renal cell carcinoma, renal pelvic carcinoma, CNS tumour, primary
CC CNS lymphoma, bone marrow tumour, brain stem nerve gliomas, pituitary
CC adenoma, or their combination. The protein is useful as a target for
CC screening new anticancer agents. The present sequence represents the 84-
CC 312 amino acid sequence of p38/JTV-1 protein.
XX
SQ Sequence 229 AA;

Query Match 64.6%; Score 1069; DB 8; Length 229;
Best Local Similarity 88.2%; Pred. No. 2.6e-101;
Matches 202; Conservative 12; Mismatches 15; Indels 0; Gaps 0;

```
Qy         84 DADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLC 143
              |||||||||:|||||||| || ||||||||||||||||||||||||||||||||||||||
Db          1 DADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLC 60

Qy        144 ERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHEYQLGFTLIWKNVPKTQMKFSVQT 203
              | :|||||||||||||:|||||:|||||| :|| | :|||||||||||||||||||:||
Db         61 EHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQDYQLGFTLIWKNVPKTQMKFSIQT 120

Qy        204 MCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIAMFQLREGSSKEKAAVFRSMNSAL 263
              |||||||||||||||||||||||||  ||||||||||:|||:|||||||||||||||||||
Db        121 MCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSAL 180

Qy        264 GRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQRWLKSCENLAFF 312
              |:|||| |||||||||||||||| || |   | |||||::||||||||
Db        181 GKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQRWMRSCENLAFF 229
```

RESULT 7
ADR86552
ID ADR86552 standard; protein; 161 AA.
XX
AC ADR86552;
XX
DT 18-NOV-2004 (first entry)
XX
DE 1-161 amino acid sequence of p38/JTV-1 protein.
XX
KW p38/JTV-1; Cytostatic; cancer; leukemia; anticancer.
XX
OS Homo sapiens.
XX
PN EP1454628-A2.
XX
PD 08-SEP-2004.
XX
PF 09-SEP-2003; 2003EP-00020344.
XX
PR 03-MAR-2003; 2003KR-00013058.
XX
PA (UYSE-) UNIV SEOUL NAT IND FOUND.
XX
PI Kim S, Park B;
XX
DR WPI; 2004-627822/61.
DR N-PSDB; ADR86549.
XX
PT New isolated p38/JTV-1 protein, useful as medicament for treating cancer
PT e.g., stomach, liver, blood, bone, pancreatic, skin, head or neck cancer
PT and cutaneous or intraocular melanoma, as well as for screening new
PT anticancer agents.
XX
PS Claim 5; SEQ ID NO 5; 47pp; English.
XX
CC The present invention relates to an isolated p38/JTV-1 protein for use as
CC medicament. The p38/JTV-1 protein or the pharmaceutical composition is
CC useful as medicament for treating breast cancer, large intestinal cancer,
CC lung cancer, small cell lung cancer, stomach cancer, liver cancer, blood
CC cancer, bone cancer, pancreatic cancer, skin cancer, head or neck cancer,
CC cutaneous or intraocular melanoma, uterine sarcoma, ovarian cancer,
CC rectal cancer, anal cancer, colon cancer, fallopian tube carcinoma,
CC endometrial carcinoma, cervical cancer, vulval cancer, vaginal carcinoma,
CC Hodgkin's disease, esophageal cancer, small intestine cancer, endocrine
CC cancer, thyroid cancer, parathyroid cancer, adrenal cancer, soft tissue
CC tumour, urethral cancer, penile cancer, prostate cancer, chronic or acute
CC leukemia, lymphocytic lymphoma, bladder cancer, kidney cancer, ureter
CC cancer, renal cell carcinoma, renal pelvic carcinoma, CNS tumour, primary
CC CNS lymphoma, bone marrow tumour, brain stem nerve gliomas, pituitary
CC adenoma, or their combination. The protein is useful as a target for
CC screening new anticancer agents. The present sequence represents the 1-
CC 161 amino acid sequence of p38/JTV-1 protein.
XX
SQ Sequence 161 AA;

Query Match 44.4%; Score 735; DB 8; Length 161;
Best Local Similarity 88.8%; Pred. No. 4.4e-67;
Matches 143; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              |||||||||||| |||||||||||||||||| ::  ||  |||||| |  ||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||||| ||||||||||:|||||||| || ||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNV 161
              |||||||||||||||||||||||| :||||||||||||:|
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSV 161
```

RESULT 8
ABP01502
ID ABP01502 standard; protein; 51 AA.
XX
AC ABP01502;
XX
DT 24-JUN-2002 (first entry)
XX

DE Human ORFX protein sequence SEQ ID NO:2986.
XX
KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypothyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW myasthenia gravis.
XX
OS Homo sapiens.
XX
PN WO200192523-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US010836.
XX
PR 30-MAY-2000; 2000US-0206132P.
PR 29-AUG-2000; 2000US-0228716P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach MD;
XX
DR WPI; 2002-106308/14.
DR N-PSDB; ABN17254.
XX
PT Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders.
XX
PS Disclosure; SEQ ID NO 2986; 1037pp; English.
XX
CC The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC in the specification). ABN15762 to ABN27252 encode the human ORFX
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage. N.B. The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 51 AA;

Query Match 15.0%; Score 248; DB 5; Length 51;
Best Local Similarity 96.1%; Pred. No. 1.3e-17;
Matches 49; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy         74 DGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALKDIVI 124
              ||||||||||||||||||||||||||||| || ||||||||||||||||||||
Db          1 DGLSKMIHTPDADLDVTNILQADEPTTLTTNALDLNSVLGKDYGALKDIVI 51
```

RESULT 9
ABB62468

ID ABB62468 standard; protein; 334 AA.
XX
AC ABB62468;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 14196.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL06571.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 14196; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 334 AA;

Query Match 13.1%; Score 216.5; DB 4; Length 334;
Best Local Similarity 26.0%; Pred. No. 4.8e-13;
Matches 86; Conservative 48; Mismatches 124; Indels 73; Gaps 9;

```
Qy          3 MYQVKPYHGGSAPLRVELPTCMYRLPNV--------HSKTTSPATDAG---------- 42
              ||::|          ::|||||| | ||           :|| :| |
Db          1 MYELKTL---LPQFDIKLPTCMYPLKNVSLAADSLASGSSTSASTSASTSSCKLEANRID 57

Qy         43 -----------------HVQETSEPSLQALESRQDDILKRLYELKAAVDGLSKMIHTPD 84
                              :|   :   :: :||: :||:| |||| :  :
Db         58 RTGRNAATCALDLDSLGRQIQRLLKDDTASVAARQEKVLKQLEELKAQLGQIR------- 110

Qy         85 ADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCE 144
              | | |    :  : ||   |        |    |:|:||| :|   | :|| |
Db        111 AGLGVCG--KTFQHTTAFQNG-------GLKEVPLQDVVINGHPNFIPYALLALKNAWRN 161

Qy        145 RYRVLSTVHTHSSVKNV-------PENLVKCFGEQARKQSRHEYQLGFTLIWKNVPKTQM 197
              | :    |||:: ::         || |      |         ::  ||||||   |:|
Db        162 LYTIDVKTFTHSTMADIGPAAREFEANLAKVPVNPALP------KISVTLIWKNCEHTEM 215

Qy        198 KFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTL---IDSWVDIAMFQLREGSSKEKAA 254
              |      || || || |:|  :   ::       |   ||  :||    ||  : | : |
Db        216 ISSPTMYVPIYGEVNIIRYLGRVGPAEYRYEGSPLCNEIDLVLDICYQLLRCNTHKTQVA 275

Qy        255 VFRSMNSALGRSPWLVGNELTVADVVLWSVL 285
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:36:20 ; Search time 17.81 Seconds
(without alignments)
1341.256 Million cell updates/sec

Title: US-10-622-817-6
Perfect score: 1655
Sequence: 1 MPMYQVKPYHGGSAPLRVEL.....RWLKSCENLAPFSTALQLLK 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1464	88.5	341	4	US-09-949-016-11312	Sequence 11312, A
2	1436	86.8	312	2	US-08-518-862C-2	Sequence 2, Appli
3	106	6.4	1512	3	US-09-443-184-48	Sequence 48, Appl
4	96.5	5.8	1440	3	US-09-357-251-37	Sequence 37, Appl
5	90.5	5.5	359	3	US-09-540-824-2	Sequence 2, Appli
6	90	5.4	559	2	US-08-756-317-10	Sequence 10, Appl
7	90	5.4	559	4	US-09-091-609-4	Sequence 4, Appli
8	90	5.4	2954	4	US-09-150-867-1	Sequence 1, Appli
9	90	5.4	5215	3	US-09-105-537-2	Sequence 2, Appli
10	89	5.4	1056	4	US-09-595-684B-29	Sequence 29, Appl
11	89	5.4	1057	3	US-09-541-782-10	Sequence 10, Appl
12	89	5.4	1057	4	US-09-723-820-10	Sequence 10, Appl
13	89	5.4	1057	4	US-10-270-085-10	Sequence 10, Appl
14	89	5.4	1287	4	US-09-949-016-7826	Sequence 7826, Ap
15	88	5.3	712	4	US-09-489-039A-10736	Sequence 10736, A
16	88	5.3	778	4	US-09-583-110-3930	Sequence 3930, Ap
17	88	5.3	786	4	US-09-107-433-3893	Sequence 3893, Ap
18	86.5	5.2	520	4	US-09-248-796A-17644	Sequence 17644, A
19	86.5	5.2	967	4	US-09-540-236-2449	Sequence 2449, Ap
20	86	5.2	535	4	US-09-489-039A-11461	Sequence 11461, A
21	85.5	5.2	556	4	US-09-134-000C-6329	Sequence 6329, Ap
22	85.5	5.2	608	4	US-09-284-768A-4	Sequence 4, Appli
23	85.5	5.2	657	4	US-09-284-768A-7	Sequence 7, Appli
24	85	5.1	222	4	US-09-248-796A-18146	Sequence 18146, A
25	85	5.1	302	4	US-09-107-532A-6924	Sequence 6924, Ap
26	85	5.1	559	4	US-09-672-749-2	Sequence 2, Appli
27	85	5.1	1971	4	US-09-914-272A-1	Sequence 1, Appli
28	85	5.1	1971	4	US-10-638-333-1	Sequence 1, Appli
29	84	5.1	443	4	US-09-949-016-10598	Sequence 10598, A
30	84	5.1	646	4	US-09-248-796A-18678	Sequence 18678, A
31	84	5.1	1057	4	US-09-428-156B-2	Sequence 2, Appli
32	83.5	5.0	360	4	US-09-949-016-10589	Sequence 10589, A
33	83.5	5.0	475	4	US-09-328-352-5379	Sequence 5379, Ap
34	83.5	5.0	515	4	US-09-489-039A-8208	Sequence 8208, Ap
35	83.5	5.0	519	4	US-09-543-681A-6505	Sequence 6505, Ap
36	83.5	5.0	789	3	US-09-002-285-84	Sequence 84, Appl
37	83.5	5.0	789	4	US-09-589-477-84	Sequence 84, Appl
38	83.5	5.0	789	4	US-10-099-285A-84	Sequence 84, Appl
39	83.5	5.0	959	4	US-09-543-681A-6879	Sequence 6879, Ap
40	83	5.0	434	3	US-09-012-072-4	Sequence 4, Appli
41	83	5.0	434	3	US-09-120-601-4	Sequence 4, Appli
42	83	5.0	448	3	US-09-120-601-6	Sequence 6, Appli
43	83	5.0	559	4	US-09-821-016-1	Sequence 1, Appli
44	83	5.0	559	4	US-10-266-787-1	Sequence 1, Appli
45	83	5.0	767	3	US-08-836-567-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-949-016-11312
; Sequence 11312, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11312
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11312

Query Match 88.5%; Score 1464; DB 4; Length 341;
Best Local Similarity 87.2%; Pred. No. 9e-153;
Matches 279; Conservative 16; Mismatches 25; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              ||||||||||| |||||||||||||||||| ::  ||  |||||| |  ||||||||||
Db         22 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 81

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||||| ||||||||||:|||||||| || ||||||||||||||||
Db         82 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 141

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              |||||||||||||||||||||| :|||||||||||||:|||||:|||||| :|| | :
Db        142 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 201

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              ||||||||||||||||||:|||||||||||||||||||||||||  ||||||||||
Db        202 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 261

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              :|||:||||||||||||||||:|||| |||||||||||||||| || |   | |||
Db        262 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 321

Qy        301 RWLKSCENLAPFSTALQLLK 320
```

```
              ||::|||||||||:|||:|||
Db        322 RWMRSCENLAPFNTALKLLK 341
```

```
RESULT 2
US-08-518-862C-2
; Sequence 2, Application US/08518862C
; Patent No. 5843757
;  GENERAL INFORMATION:
;    APPLICANT:  Vogelstein, Bert
;    APPLICANT:  Kinzler, Kenneth W.
;    APPLICANT:  Nicolaides, Nicholas C.
;    TITLE OF INVENTION:  Human JTV1 Gene Overlaps PMS2 Gene
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Witcoff, Ltd.
;      STREET:  1001 G Street, N.W.
;      CITY:  Washington, D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/518,862C
;      FILING DATE:  24-AUG-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kagan, Sarah A.
;      REGISTRATION NUMBER:  32,141
;      REFERENCE/DOCKET NUMBER:  01107.49697
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-508-9100
;      TELEFAX:  202-508-9299
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  312 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-518-862C-2

  Query Match             86.8%;  Score 1436;  DB 2;  Length 312;
  Best Local Similarity   87.5%;  Pred. No. 9.6e-150;
  Matches  273;  Conservative   14;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              |||||||||||| |||||||||||||||| ::  ||  |||||| |  ||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||| ||||||||||:|||||||| || ||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              |||||||||||||||||||||| :|||||||||||||:|||||:|||||| :|| | :
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              ||||||||||||||||||:|||||||||||||||||||||||||  ||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              :|||:||||||||||||||||:|||| |||||||||||||| || |   | |||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWLKSCENLAPF 312
              ||::|||||||
Db        301 RWMRSCENLAPF 312
```

```
RESULT 3
US-09-443-184-48
; Sequence 48, Application US/09443184A
; Patent No. 6372431
; GENERAL INFORMATION:
;  APPLICANT: Cunningham, Mary Jane
;  APPLICANT: Zweiger, Gary
;  APPLICANT: Kaser, Matthew R.
;  APPLICANT: Panzer, Scott
;  APPLICANT: Seilhammer, Jeffrey J.
;  APPLICANT: Yue, Henry
;  APPLICANT: Baughn, Mariah
;  APPLICANT: Azimzai, Yalda
;  APPLICANT: Lal, Preeti
;  TITLE OF INVENTION: MAMMALIAN TOXICOLOGICAL RESPONSE MARKERS
;  FILE REFERENCE: PC-0007 US
;  CURRENT APPLICATION NUMBER: US/09/443,184A
;  CURRENT FILING DATE: 1999-11-19
;  NUMBER OF SEQ ID NOS: 138
;  SOFTWARE: PERL Program
; SEQ ID NO 48
;   LENGTH: 1512
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6372431 2302721CD1
US-09-443-184-48

  Query Match             6.4%;  Score 106;  DB 3;  Length 1512;
  Best Local Similarity   21.8%;  Pred. No. 0.096;
  Matches   44;  Conservative   33;  Mismatches   69;  Indels   56;  Gaps    8;

Qy        124 INANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHEYQL 183
              :  |   |||  |:                    :|::| :: |    |: ::   |
Db          6 LTVNSGDPPLGALL------------------AVEHVKDD-VSISVEEGKENILH---- 41

Qy        184 GFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFL------FSLFGQKHNAVTLTLIDSWV 237
              : :||  | :                  :| |:|       |:|   | :  | || |:
Db         42 ----VSENVIFTDV-------------NSILRYLARVATTAGLYGS--NLMEHTEIDHWL 82

Qy        238 DIAMFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQ-----QTGGSS 292
              : :  :|   |     :    :|  |    :|||| |::||: :|: |:     |
Db         83 EFSATKL--SSCDSFTSTINELNHCLSLRTYLVGNSLSLADLCVWATLKGNAAWQEQLKQ 140

Qy        293 GAAPTNVQRWLKSCENLAPFST 314
              || :|:||    |    | :
Db        141 KKAPVHVKRWFGFLEAQQAFQS 162
```

```
RESULT 4
US-09-357-251-37
; Sequence 37, Application US/09357251
; Patent No. 6271441
; GENERAL INFORMATION:
;  APPLICANT: Falco, S. Carl
;  APPLICANT: Famodu, Layo O.
;  APPLICANT: Orozco, Buddy
;  APPLICANT: Schwaber, James S.
;  TITLE OF INVENTION: Plant Aminoacyl-tRNA Synthetase
;  FILE REFERENCE: BB-1193
;  CURRENT APPLICATION NUMBER: US/09/357,251
;  CURRENT FILING DATE: 1999-07-20
;  EARLIER APPLICATION NUMBER: 60/093,530
;  EARLIER FILING DATE: July 21, 1998
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 37
;   LENGTH: 1440
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-357-251-37
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:57:26 ; Search time 50.2443 Seconds
(without alignments)
2084.158 Million cell updates/sec

Title: US-10-622-817-6
Perfect score: 1655
Sequence: 1 MPMYQVKPYHGGSAPLRVEL.....RWLKSCENLAPFSTALQLLK 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1380268 seqs, 327241040 residues

Total number of hits satisfying chosen parameters: 1380268

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1436	86.8	312	16	US-10-463-676-4	Sequence 4, Appli
2	1069	64.6	229	16	US-10-463-676-6	Sequence 6, Appli
3	735	44.4	161	16	US-10-463-676-5	Sequence 5, Appli
4	105.5	6.4	925	16	US-10-437-963-126132	Sequence 126132,
5	104	6.3	980	15	US-10-369-493-6336	Sequence 6336, Ap
6	103.5	6.3	468	15	US-10-424-599-260388	Sequence 260388,
7	99.5	6.0	914	16	US-10-437-963-153870	Sequence 153870,
8	99	6.0	691	15	US-10-282-122A-44369	Sequence 44369, A
9	98	5.9	690	9	US-09-815-242-5841	Sequence 5841, Ap
10	96.5	5.8	1440	10	US-09-864-464-37	Sequence 37, Appl
11	96.5	5.8	1440	16	US-10-786-720-31	Sequence 31, Appl
12	96	5.8	569	16	US-10-437-963-123904	Sequence 123904,
13	95	5.7	722	15	US-10-369-493-10013	Sequence 10013, A
14	93.5	5.6	456	15	US-10-425-114-63166	Sequence 63166, A
15	93	5.6	204	15	US-10-369-493-10805	Sequence 10805, A
16	93	5.6	1398	16	US-10-437-963-180124	Sequence 180124,
17	91	5.5	201	15	US-10-369-493-13717	Sequence 13717, A
18	91	5.5	455	15	US-10-425-114-42633	Sequence 42633, A
19	91	5.5	805	15	US-10-424-599-191796	Sequence 191796,
20	91	5.5	1319	16	US-10-437-963-197783	Sequence 197783,
21	91	5.5	1788	16	US-10-437-963-197780	Sequence 197780,
22	90.5	5.5	419	15	US-10-424-599-260387	Sequence 260387,
23	90.5	5.5	578	15	US-10-418-861B-32	Sequence 32, Appl
24	90.5	5.5	661	15	US-10-369-493-2558	Sequence 2558, Ap
25	90.5	5.5	716	15	US-10-369-493-2175	Sequence 2175, Ap
26	90	5.4	559	9	US-09-364-847-21	Sequence 21, Appl
27	90	5.4	856	9	US-09-364-847-33	Sequence 33, Appl
28	90	5.4	856	9	US-09-364-847-35	Sequence 35, Appl
29	90	5.4	1500	16	US-10-437-963-111265	Sequence 111265,
30	90	5.4	5215	9	US-09-861-289-2	Sequence 2, Appli
31	90	5.4	5215	9	US-09-860-846-2	Sequence 2, Appli
32	90	5.4	5215	10	US-09-988-384B-2	Sequence 2, Appli
33	90	5.4	5215	10	US-09-836-821-2	Sequence 2, Appli
34	90	5.4	5215	14	US-10-271-889-45	Sequence 45, Appl
35	89.5	5.4	203	15	US-10-369-493-8604	Sequence 8604, Ap
36	89	5.4	407	15	US-10-424-599-267113	Sequence 267113,
37	89	5.4	1056	15	US-10-282-174-472	Sequence 472, App
38	89	5.4	1056	15	US-10-282-174-474	Sequence 474, App
39	89	5.4	1056	15	US-10-282-174-476	Sequence 476, App
40	89	5.4	1056	17	US-10-600-009-472	Sequence 472, App
41	89	5.4	1056	17	US-10-600-009-474	Sequence 474, App
42	89	5.4	1056	17	US-10-600-009-476	Sequence 476, App
43	89	5.4	1057	16	US-10-714-796-4	Sequence 4, Appli
44	88.5	5.3	656	15	US-10-369-493-18559	Sequence 18559, A
45	88.5	5.3	799	16	US-10-437-963-112603	Sequence 112603,

ALIGNMENTS

RESULT 1
US-10-463-676-4
; Sequence 4, Application US/10463676
; Publication No. US20040175375A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Sunghoon
; APPLICANT: Park, Bum-Joon
; TITLE OF INVENTION: Method for Treating Cancer Using P38/JTV-1 and Method
; TITLE OF INVENTION: for Screening Pharmaceutical Composition for Treating Cancer
; FILE REFERENCE: 012679-091
; CURRENT APPLICATION NUMBER: US/10/463,676
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: KR 10-2003-13058
; PRIOR FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)...(312)
; OTHER INFORMATION: 1-312 amino acid sequence of p38/JTV-1
US-10-463-676-4

Query Match 86.8%; Score 1436; DB 16; Length 312;
Best Local Similarity 87.5%; Pred. No. 2.6e-137;
Matches 273; Conservative 14; Mismatches 25; Indels 0; Gaps 0;

```
Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              |||||||||||| |||||||||||||||| ::  ||  |||||| |  ||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||||| ||||||||||:|||||||| || ||||||||||||||||
```

```
Db          61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy         121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
               |||||||||||||||||||||||| :|||||||||||||:|||||:|||||| :|| | :
Db         121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy         181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
               ||||||||||||||||||:||||||||||||||||||||||||||  ||||||||||
Db         181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy         241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
               :|||:||||||||||||||||:|||| |||||||||||||||| || |    | |||
Db         241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy         301 RWLKSCENLAPF 312
               ||::|||||||| 
Db         301 RWMRSCENLAPF 312


RESULT 2
US-10-463-676-6
; Sequence 6, Application US/10463676
; Publication No. US20040175375A1
; GENERAL INFORMATION:
;  APPLICANT: Kim, Sunghoon
;  APPLICANT:  Park, Bum-Joon
;  TITLE OF INVENTION: Method for Treating Cancer Using P38/JTV-1 and Method
;  TITLE OF INVENTION:  for Screening Pharmaceutical Composition for Treating Cancer
;  FILE REFERENCE: 012679-091
;  CURRENT APPLICATION NUMBER: US/10/463,676
;  CURRENT FILING DATE:  2003-06-18
;  PRIOR APPLICATION NUMBER: KR 10-2003-13058
;  PRIOR FILING DATE: 2003-03-03
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 229
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)...(229)
;   OTHER INFORMATION: 84-312 amino acid sequence p38/JTV-1
US-10-463-676-6

  Query Match           64.6%;  Score 1069;  DB 16;  Length 229;
  Best Local Similarity   88.2%;  Pred. No. 4e-100;
  Matches  202;  Conservative   12;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          84 DADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLC 143
               |||||||||:|||||||| || ||||||||||||||||||||||||||||||||||||||
Db           1 DADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLC 60

Qy         144 ERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHEYQLGFTLIWKNVPKTQMKFSVQT 203
               | :|||||||||||||:|||||:|||||| :|| | :|||||||||||||||||||:||
Db          61 EHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQDYQLGFTLIWKNVPKTQMKFSIQT 120

Qy         204 MCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIAMFQLREGSSKEKAAVFRSMNSAL 263
               |||||||||||||||||||||||  ||||||||||:|||:|||||||||||||||||||
Db         121 MCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIAIFQLKEGSSKEKAAVFRSMNSAL 180

Qy         264 GRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQRWLKSCENLAPF 312
               |:|||| |||||||||||||||| || |    | |||||::||||||||
Db         181 GKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQRWMRSCENLAPF 229


RESULT 3
US-10-463-676-5
; Sequence 5, Application US/10463676
; Publication No. US20040175375A1
; GENERAL INFORMATION:
;  APPLICANT: Kim, Sunghoon
;  APPLICANT:  Park, Bum-Joon
;  TITLE OF INVENTION: Method for Treating Cancer Using P38/JTV-1 and Method
;  TITLE OF INVENTION:  for Screening Pharmaceutical Composition for Treating Cancer
;  FILE REFERENCE: 012679-091
;  CURRENT APPLICATION NUMBER: US/10/463,676
;  CURRENT FILING DATE:  2003-06-18
;  PRIOR APPLICATION NUMBER: KR 10-2003-13058
;  PRIOR FILING DATE: 2003-03-03
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 161
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)...(161)
;   OTHER INFORMATION: 1-161 amino acid sequence of p38/JTV-1
US-10-463-676-5

  Query Match           44.4%;  Score 735;  DB 16;  Length 161;
  Best Local Similarity   88.8%;  Pred. No. 2.5e-66;
  Matches  143;  Conservative    5;  Mismatches   13;  Indels    0;  Gaps    0;

Qy           1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
               |||||||||||| |||||||||||||||||| ::  ||  |||||| |  ||||||||||
Db           1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy          61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
               |||||||||||||||||||| ||||||||||:|||||||| || ||||||||||||||||
Db          61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy         121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNV 161
               |||||||||||||||||||||||| :|||||||||||||:|
Db         121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSV 161


RESULT 4
US-10-437-963-126132
; Sequence 126132, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 126132
;   LENGTH: 925
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_28709C.1.pep
US-10-437-963-126132

  Query Match           6.4%;  Score 105.5;  DB 16;  Length 925;
  Best Local Similarity   20.3%;  Pred. No. 0.51;
  Matches   64;  Conservative   46;  Mismatches  107;  Indels   99;  Gaps   13;

Qy          25 YRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQDDI--LKRLYELKAAVDGLSKMIHT 82
               | : |: :| | |          | :| ||||  :  ::  ::: |    :  |   :|
Db         152 YMVDNIIAKVTGPV--------TIDPRLQALYKKTTELVGIEKQSEKLVKILSLGDDVHA 203

Qy          83 PDADLDVTNILQAD--EPTTLATNTLDLNSVLGKDYGA------------LKDIVINAN 127
               |  : : :|:       |||:    | :  |  | ||             |:||:|: :
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:34:50 ; Search time 13.3213 Seconds
(without alignments)
2311.294 Million cell updates/sec

Title: US-10-622-817-6
Perfect score: 1655
Sequence: 1 MPMYQVKPYHGGSAPLRVEL.....RWLKSCENLAPFSTALQLLK 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	126.5	7.6	719	2	T52043	probable glutamate
2	112	6.8	728	2	T01200	probable glutamate
3	104	6.3	980	2	T30089	probable zinc prot
4	99	6.0	691	2	B89797	glycerol ester hyd
5	98.5	6.0	222	2	G82441	probable glutathio
6	98.5	6.0	377	2	F69172	hypothetical prote
7	98	5.9	690	2	A24545	triacylglycerol li
8	96.5	5.8	1440	1	SYHUQT	multifunctional am
9	95.5	5.8	753	2	T24869	hypothetical prote
10	95	5.7	488	1	H64313	corrinoid/iron-sul
11	95	5.7	552	2	D82421	conserved hypothet
12	95	5.7	2655	2	D96595	probable acetyl-Co
13	94.5	5.7	2471	2	T42977	large tegument pro
14	94	5.7	435	2	T01567	hypothetical prote
15	94	5.7	436	2	C86719	GTP-binding protei
16	93.5	5.6	1064	2	B86465	probable Protein k
17	93	5.6	490	2	D84999	ketol-acid reducto
18	92	5.6	754	2	S37403	transcription fact
19	91.5	5.5	436	2	I51237	translation elonga
20	91	5.5	1265	2	T47626	structural mainten
21	90.5	5.5	359	2	T37921	ras-associated pro
22	90.5	5.5	456	2	C84764	hypothetical prote
23	90.5	5.5	642	2	T39490	transcription init
24	90.5	5.5	716	2	T37830	probable glutamate
25	90	5.4	437	2	I51238	translation elonga
26	90	5.4	559	2	A38604	poly(3-hydroxyalka
27	90	5.4	2954	2	T14156	kinesin-related pr
28	89.5	5.4	317	2	T05528	hypothetical prote
29	89.5	5.4	1597	2	S65053	genome polyprotein
30	89.5	5.4	1714	1	S18644	multifunctional am
31	89.5	5.4	4976	2	T14165	peptide synthetase
32	89	5.4	184	2	S74356	glutathione S-tran
33	89	5.4	763	2	AE2443	penicillin-binding
34	89	5.4	1023	2	T31669	neural zinc finger
35	89	5.4	1056	1	G02157	kinesin-like spind
36	89	5.4	1247	2	A33812	interphotoreceptor
37	88.5	5.3	602	2	AB3542	gtp-binding protei
38	88.5	5.3	656	2	H84206	acetyl-CoA synthet
39	88	5.3	821	2	T24728	hypothetical prote
40	88	5.3	2279	2	T42531	acetyl-CoA carboxy
41	88	5.3	2280	2	T38906	acetyl-CoA carboxy
42	87.5	5.3	474	2	B70597	hypothetical prote
43	87.5	5.3	820	2	H82302	ATP-dependent heli
44	87	5.3	570	2	T15763	hypothetical prote
45	86	5.2	906	2	T45158	pre-mRNA splicing

ALIGNMENTS

RESULT 1
T52043
probable glutamate-tRNA ligase (EC 6.1.1.17) [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004
C;Accession: T52043
R;Day, I.S.; Golovkin, M.; Reddy, A.S.
Biochim. Biophys. Acta 1399, 219-24, 1998
A;Title: Cloning of the cDNA for glutamyl-tRNA synthetase from Arabidopsis thaliana.
A;Reference number: Z24836; MUID:9765600; PMID:9765600
A;Accession: T52043
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-719 <DAY>
A;Cross-references: UNIPROT:O82462; EMBL:AF067773; PIDN:AAC36469.1
C;Superfamily: yeast glutamate-tRNA ligase; glutamine-tRNA ligase homology
C;Keywords: ligase

Query Match 7.6%; Score 126.5; DB 2; Length 719;
Best Local Similarity 28.0%; Pred. No. 0.017;
Matches 49; Conservative 22; Mismatches 57; Indels 47; Gaps 8;

```
Qy       128 PASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHEYQLGFTL 187
            | |||||::|  |     |  |: : ::   ||  :      || :    | |:
Db        10 PESPPLSVIVALSLSASPV----TIDSSAAATTVPSFVF----SDGRKLN------GATV 55

Qy       188 IWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIAMFQLREG 247
               : : |                  |  |: |   :|  :||   : || ||| |      |
Db        56 LLRYV-----------------GRSAKKLPDFYG--NNAFDSSQIDEWVDYASV-FSSG 94

Qy       248 SSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQRW 302
               |  | |      ::  |  | :|||: |::||| :|| |  ||          |||
Db        95 SEFENAC--GRVDKYLESSTFLVGHSLSIADVAIWSALAGTG----------QRW 137
```

RESULT 2
T01200
probable glutamate-tRNA ligase (EC 6.1.1.17) F21E10.12 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: T01200
R;Davidson, S.; Rohlfing, T.; David, M.; O'Brian, D.
submitted to the EMBL Data Library, April 1998
A;Description: The sequence of A. thaliana F21E10.
A;Reference number: Z14258
A;Accession: T01200
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-728 <DAV>
A;Cross-references: UNIPROT:O65253; EMBL:AF058914; NID:g3047074; PID:g3047084; GSPDB:GN0
A;Experimental source: cultivar Columbia
C;Genetics:

A;Gene: ATSP:F21E10.12
A;Map position: 5
A;Introns: 47/2; 89/3; 141/1; 503/3; 659/3
C;Superfamily: yeast glutamate-tRNA ligase; glutamine-tRNA ligase homology
C;Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis
F;223-499/Domain: glutamine-tRNA ligase homology <EGL>

Query Match 6.8%; Score 112; DB 2; Length 728;
Best Local Similarity 26.6%; Pred. No. 0.29;
Matches 49; Conservative 22; Mismatches 57; Indels 56; Gaps 9;

```
Qy        128 PASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHEYQLGFTL 187
              |  |||||::|   |         |: : ::   ||  :        || :      | |:
Db         10 PESPPLSVIVALSLSASPV----TIDSSAAATTVPSFVF----SDGRKLN------GATV 55

Qy        188 IWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLT---------LIDSWVD 238
              : : |                |  |: |   :|  :||   :           || |||
Db         56 LLRYV-----------------GRSAKKLPDFYG--NNAFDSSQVSILCINMKIDEWVD 95

Qy        239 IAMFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTN 298
               |      ||  | |     ::  |  | :|||: |::||| :|| |  ||
Db         96 YASV-FSSGSEFENAC--GRVDKYLESSTFLVGHSLSIADVAIWSALAGTG--------- 143

Qy        299 VQRW 302
               |||
Db        144 -QRW 146
```

RESULT 3
T30089
probable zinc proteinase (EC 3.4.24.-) C02G6.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T30089
R;Bentley, D.; Kemp, K.; Scheet, P.
submitted to the EMBL Data Library, April 1996
A;Description: The sequence of C. elegans cosmid C02G6.
A;Reference number: Z20734
A;Accession: T30089
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-980 <BEN>
A;Cross-references: UNIPROT:Q17592; EMBL:U55372; PIDN:AAA98001.1; GSPDB:GN00023; CESP:C0
A;Experimental source: strain Bristol N2; clone C02G6
C;Genetics:
A;Gene: CESP:C02G6.1
A;Map position: 5
A;Introns: 25/3; 215/2; 266/3; 540/3; 585/3; 786/1; 898/2
C;Superfamily: insulin-degrading enzyme (IDE)
C;Keywords: hydrolase; metalloproteinase; zinc
F;70,74/Binding site: zinc (His) #status predicted
F;73/Active site: Glu #status predicted

Query Match 6.3%; Score 104; DB 2; Length 980;
Best Local Similarity 20.1%; Pred. No. 2.1;
Matches 86; Conservative 67; Mismatches 133; Indels 142; Gaps 22;

```
Qy         10 HGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQDDILKRLYE- 68
              | |    | |||    : : :: | : : |   | :  | : |   ||:  |:|:: :
Db        276 HKGPGSLLVELKRLGW-VNSLKSDSNTIAAGFGILNVTMDLSTGGLEN-VDEIIQLMLNY 333

Qy         69 ---LKAAVDGLSKMIHTPDADL-DVT-NILQADEPTTLATNT------LDLNSVLGKDY- 116
                ||:   |  : ||   ||| ||        ::|  :| |       : :  :|   |
Db        334 IGMLKSF--GPQQWIHDELADLSDVKFRFKDKEQPMKMAINIAASLQYIPIEHILSSRYL 391

Qy        117 ------GALKDIVINANPASPPLSLLVLHRLLCERYR---------------------- 147
                   :|:::    |::      :| |:: ::::
Db        392 LTKYEPERIKELLSTLTPSN------MLVRVVSQKFKEQEGNTNEPVYGTEMKVTDISPE 445

Qy        148 -------VLSTVH--THSSVKNVPENLVKCFGEQARKQSRHEYQL-----GFTLIW---- 189
                    | | |   |  ||  | :   ||:: |:  ::|:       |:: :|
Db        446 KMKKYENALKTSHHALHLPEKN--EYIATNFGQKPRESVKNEHPKLISDDGWSRVWFKQD 503
```

```
Qy        190 --KNVPKTQMKFSVQT------------------MCPIEGE--------GNIARFLFSL 220
                |:|| : ||:: |                  | |  |         |   :|  |
Db        504 DEYNMPKQETKFALTTPIVSQNPRISLISSLWLWCFCDILSEETYNAALAGLGCQFELSP 563

Qy        221 FG-----------QKHNAVTL-----------------TLIDSWVDIAMFQLREGSSKE 251
              ||            ::| ::||                  :|:  :|   |::   | |
Db        564 FGVQKQSTDGREAERHASLTLHVYGYDEKQPLFVKHLTSCMINFKIDRTRFEVLFESLK- 622

Qy        252 KAAVFRSMNSALGRSPWLVG---NELTVADVVLWSVLQQTGGSSGAAPTNVQRWLKSCEN 308
                |:: :     |:|:    |:| : | | ||  |             ||| : :  |
Db        623 -----RTLTNHAFSQPYLLTQHYNQLLIVDKV-WSKEQLLAVCDSVTLENVQGFAR--EM 674

Qy        309 LAPFSTAL 316
              |  |  |
Db        675 LQAFHMEL 682
```

RESULT 4
B89797
glycerol ester hydrolase [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: B89797
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of meticillin-resistant Stapylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: B89797
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-691 <KUR>
A;Cross-references: UNIPROT:Q99WQ6; GB:BA000018; PID:g13700235; PIDN:BAB41533.1; GSPDB:G
A;Experimental source: strain N315
C;Genetics:
A;Gene: geh
C;Superfamily: Staphylococcus triacylglycerol lipase

Query Match 6.0%; Score 99; DB 2; Length 691;
Best Local Similarity 20.2%; Pred. No. 3.4;
Matches 65; Conservative 53; Mismatches 110; Indels 94; Gaps 15;

```
Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDA--------GHVQETSEPSL 52
              :|  :| | |  :||      |:          : ||    :| |         : |:| :|:
Db        172 IPAQKVLPNHDKAAPTSTTPPS--------NDKTAPKSTKAQDATTDKHPNQQDTHQPAH 223

Qy         53 QALESRQDDILKRLYELKAAVDGLSKMI---HTPDADLD--------VTNILQADEPTTL 101
              | ::::||| : |  | |  |  ||| |   ::|:   |         : : |||  | |
Db        224 QIIDAKQDDTV-RQSEQKPQVGDLSKHIDGQNSPEKPTDKNTDNKQLIKDALQA--PKTR 280

Qy        102 ATNTLDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNV 161
              :|     :         | |   : ||    |              :| |    |  |  :  |
Db        281 STTNAAAD--------AKKVRPLKANQVQP----------LNKYPV---VFVHGFLGLV 318

Qy        162 PENLVKCFG----------EQARKQSRHEYQLGFTLIWKNVPK-TQMKFSVQTMCPIEG 209
              :|    :           |: |||  : :|   :    |  :  :: : ::      |
Db        319 GDNAPALYPNYWGGNKFKVIEELRKQGYNVHQASVSAFGSNYDRAVELYYYIKGGRVDYG 378

Qy        210 EGNIARFLFSLFGQKHNAV-------------------TLTLIDSWVDIAMFQLREGSS 249
              : |::    :|: :  :                    |: |:: :       || | :
Db        379 AAHAAKYGHERYGKTYKGIMPNWEPGKKVHLVGHSMGGQTIRLMEEF-------LRNG-N 430

Qy        250 KEKAAVFRSMNSALGRSPWLVG 271
              ||: |  ::    :  ||   |
Db        431 KEEIAYHKAHGGEI--SPLFTG 450
```

RESULT 5
G82441
probable glutathione S-transferase VCA0584 [imported] - Vibrio cholerae (strain N16961 s
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:33:04 ; Search time 60.8145 Seconds
(without alignments)
2694.512 Million cell updates/sec

Title: US-10-622-817-6
Perfect score: 1655
Sequence: 1 MPMYQVKPYHGGSAPLRVEL.....RWLKSCENLAPFSTALQLLK 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1655	100.0	320	2	Q8R010	Q8r010 mus musculu
2	1652	99.8	320	2	Q8R2Y6	Q8r2y6 mus musculu
3	1573	95.0	320	1	MCA2_CRIGR	Q9wvm7 cricetulus
4	1464	88.5	320	1	MCA2_HUMAN	Q13155 homo sapien
5	1460	88.2	320	2	Q96CZ5	Q96cz5 homo sapien
6	1412	85.3	280	2	Q8R3V2	Q8r3v2 mus musculu
7	1153.5	69.7	311	2	Q6DK86	Q6dk86 xenopus tro
8	1146.5	69.3	311	2	Q6INU4	Q6inu4 xenopus lae
9	1146.5	69.3	311	2	Q7ZYD7	Q7zyd7 xenopus lae
10	926.5	56.0	321	2	Q7T3C0	Q7t3c0 brachydanio
11	247	14.9	340	2	Q7Q7A3	Q7q7a3 anopheles g
12	224	13.5	301	2	Q7KUM5	Q7kum5 drosophila
13	224	13.5	313	2	Q8T060	Q8t060 drosophila
14	224	13.5	322	2	Q6NKM4	Q6nkm4 drosophila
15	216.5	13.1	334	1	MCA2_DROME	Q9vur3 drosophila
16	126.5	7.6	719	2	O82462	O82462 arabidopsis
17	112	6.8	728	2	O65253	O65253 arabidopsis
18	107	6.5	715	2	Q9LIZ8	Q9liz8 oryza sativ
19	106	6.4	328	2	Q86X73	Q86x73 homo sapien
20	106	6.4	559	2	Q8RPZ6	Q8rpz6 pseudomonas
21	106	6.4	869	2	Q6PD57	Q6pd57 homo sapien
22	106	6.4	1721	2	Q7NID9	Q7nid9 gloeobacter
23	104	6.3	980	2	Q17592	Q17592 caenorhabdi
24	102.5	6.2	420	2	Q8FQQ0	Q8fqq0 corynebacte
25	102	6.2	559	2	Q939A8	Q939a8 pseudomonas
26	102	6.2	1048	2	Q7N3E2	Q7n3e2 photorhabdu
27	101.5	6.1	715	2	Q91B77	Q91b77 white spot
28	101.5	6.1	730	2	Q910F5	Q910f5 white spot
29	101.5	6.1	913	2	Q6DRB3	Q6drb3 brachydanio
30	101.5	6.1	913	2	Q6PFQ4	Q6pfq4 brachydanio
31	100.5	6.1	555	2	Q7MEV7	Q7mev7 vibrio vuln
32	100.5	6.1	555	2	Q8D7U3	Q8d7u3 vibrio vuln
33	100.5	6.1	885	2	Q8K2J4	Q8k2j4 mus musculu
34	99.5	6.0	922	2	Q6TEM5	Q6tem5 brachydanio
35	99	6.0	559	2	Q8RQ67	Q8rq67 pseudomonas
36	99	6.0	690	1	LIP_STAAW	Q8nyc2 staphylococ
37	99	6.0	690	2	Q6GCF1	Q6gcf1 staphylococ
38	99	6.0	691	2	Q79SZ7	Q79sz7 staphylococ
39	99	6.0	691	2	Q99WQ6	Q99wq6 staphylococ
40	99	6.0	691	2	Q7A7P2	Q7a7p2 staphylococ
41	98.5	6.0	222	2	Q9KM05	Q9km05 vibrio chol
42	98.5	6.0	377	2	O26161	O26161 methanobact
43	98.5	6.0	1597	2	Q8V9A2	Q8v9a2 crucifer to
44	98	5.9	240	2	Q8BTR1	Q8btr1 mus musculu
45	98	5.9	690	1	LIP_STAAU	P10335 staphylococ

ALIGNMENTS

RESULT 1
Q8R010
ID Q8R010 PRELIMINARY; PRT; 320 AA.
AC Q8R010;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Jtv1-pending protein.
GN Name=Jtv1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N, and FVB/N-3; TISSUE=Mammary tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N-3; TISSUE=Mammary tumor;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC024410; AAH24410.1; -.
DR EMBL; BC026972; AAH26972.1; -.
DR MGD; MGI:2385237; Jtv1.
DR InterPro; IPR004046; GST_Cterm.
DR InterPro; IPR010987; GST_C_like.
DR Pfam; PF00043; GST_C; 1.
SQ SEQUENCE 320 AA; 35396 MW; 1745D7EE4BC3670D CRC64;

```
  Query Match           100.0%;  Score 1655;  DB 2;  Length 320;
  Best Local Similarity  100.0%;  Pred. No. 2.6e-128;
  Matches  320;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300

Qy        301 RWLKSCENLAPFSTALQLLK 320
              ||||||||||||||||||||
Db        301 RWLKSCENLAPFSTALQLLK 320


RESULT 2
Q8R2Y6
ID   Q8R2Y6      PRELIMINARY;      PRT;   320 AA.
AC   Q8R2Y6;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Jtv1-pending protein.
GN   Name=Jtv1;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CZECH II; TISSUE=Mammary tumor;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CZECH II; TISSUE=Mammary tumor;
RA   Strausberg R.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC026958; AAH26958.1; -.
DR   MGD; MGI:2385237; Jtv1.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   320 AA;  35423 MW;  1C21F1A74C9882B4 CRC64;

  Query Match           99.8%;  Score 1652;  DB 2;  Length 320;
  Best Local Similarity  99.7%;  Pred. No. 4.6e-128;
  Matches  319;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              |||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTNPATDAGHVQETSEPSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300

Qy        301 RWLKSCENLAPFSTALQLLK 320
              ||||||||||||||||||||
Db        301 RWLKSCENLAPFSTALQLLK 320


RESULT 3
MCA2_CRIGR
ID   MCA2_CRIGR     STANDARD;      PRT;   320 AA.
AC   Q9WVM7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Multisynthetase complex auxiliary component p38.
OS   Cricetulus griseus (Chinese hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Cricetulus.
OX   NCBI_TaxID=10029;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=99096915; PubMed=9878398; DOI=10.1006/jmbi.1998.2316;
RA   Quevillon S., Robinson J.-C., Berthonneau E., Siatecka M., Mirande M.;
RT   "Macromolecular assemblage of aminoacyl-tRNA synthetases:
RT   identification of protein-protein interactions and characterization of
RT   a core protein.";
RL   J. Mol. Biol. 285:183-195(1999).
CC   -!- FUNCTION: Probable core protein of the multisynthetase complex
CC       that serves as a template for the assembly of the supramolecular
CC       structure.
CC   -!- SUBUNIT: Component of the multisynthetase complex which is
CC       comprised of a bifunctional glutamyl-prolyl-tRNA synthetase, the
CC       monospecific isoleucyl, leucyl, glutaminyl, methionyl, lysyl,
CC       arginyl, and aspartyl-tRNA synthetases as well as three auxiliary
CC       proteins, p18, p48 and p43.
CC   -!- SIMILARITY: Contains 1 GST-like domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF072727; AAD38422.1; -.
DR   InterPro; IPR010987; GST_C_like.
DR   InterPro; IPR004046; GST_Cterm.
DR   Pfam; PF00043; GST_C; 1.
KW   Protein biosynthesis.
SQ   SEQUENCE   320 AA;  35433 MW;  6D24E033ABEC810A CRC64;

  Query Match             95.0%;  Score 1573;  DB 1;  Length 320;
  Best Local Similarity   94.4%;  Pred. No. 1.6e-121;
  Matches  302;  Conservative    6;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPMYQVKSYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||||||||||||||||||||  | || |||| ||||||||||
Db         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPVPLTTNALDLNLVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              ||||||||||||:||||||||||||||||||||||||||||||:|||||||||||||||
Db        121 DIVINANPASPPISLLVLHRLLCERYRVLSTVHTHSSVKNVPENLLKCFGEQARKQSRHE 180

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              ||||||||||||||||||||||||||||||||||||||||||||:|| ||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHSAVNLTLIDSWVDIA 240

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              ||||:||||||:||||||||||||:||||||||||||||||||||||  || || |||
Db        241 MFQLKEGSSKERAAVFRSMNSALGKSPWLVGNELTVADVVLWSVLQQTGDGSGVAPANVQ 300

Qy        301 RWLKSCENLAPFSTALQLLK 320
              ||||||||| ||||||||||
Db        301 RWLKSCENLVPFSTALQLLK 320


RESULT 4
MCA2_HUMAN
ID   MCA2_HUMAN     STANDARD;      PRT;   320 AA.
AC   Q13155; Q9P1L2;
DT   01-NOV-1997 (Rel. 35, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Multisynthetase complex auxiliary component p38 (JTV-1 protein)
DE   (PRO0992).
GN   Name=JTV1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96115582; PubMed=8666379;
RA   Nicolaides N.C., Kinzler K.W., Vogelstein B.;
RT   "Analysis of the 5' region of PMS2 reveals heterogeneous transcripts
RT   and a novel overlapping gene.";
RL   Genomics 29:329-334(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lymph;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   SEQUENCE OF 197-320 FROM N.A.
RC   TISSUE=Fetal liver;
RA   Zhang C., Yu Y., Zhang S., Wei H., Zhou G., Ouyang S., Luo L., Bi J.,
RA   Liu M., He F.;
RT   "Functional prediction of the coding sequences of 121 new genes
RT   deduced by analysis of cDNA clones from human fetal liver.";
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   INTERACTION WITH FUBP1.
RX   MEDLINE=22716800; PubMed=12819782; DOI=10.1038/ng1182;
RA   Kim M.J., Park B.-J., Kang Y.-S., Kim H.J., Park J.-H., Kang J.W.,
RA   Lee S.W., Han J.M., Lee H.-W., Kim S.;
RT   "Downregulation of FUSE-binding protein and c-myc by tRNA synthetase
RT   cofactor p38 is required for lung cell differentiation.";
RL   Nat. Genet. 34:330-336(2003).
CC   -!- FUNCTION: Probable core protein of the multisynthetase complex
CC       that serves as a template for the assembly of the supramolecular
CC       structure. Mediates ubiquitination of FUBP1 and its degradation by
CC       the proteasome.
CC   -!- SUBUNIT: Component of the multisynthetase complex which is
CC       comprised of a bifunctional glutamyl-prolyl-tRNA synthetase, the
CC       monospecific isoleucyl, leucyl, glutaminyl, methionyl, lysyl,
CC       arginyl, and aspartyl-tRNA synthetases as well as three auxiliary
CC       proteins, p18, p48 and p43. Binds FUBP1.
CC   -!- SIMILARITY: Contains 1 GST-like domain.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC       frameshift in position 312.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U24169; AAC50391.1; ALT_FRAME.
DR   EMBL; BC002853; AAH02853.1; -.
DR   EMBL; BC010156; AAH10156.1; -.
DR   EMBL; AF116615; AAF71039.1; -.
DR   H-InvDB; HIX0006460; -.
DR   MIM; 600859; -.
DR   InterPro; IPR010987; GST_C_like.
DR   InterPro; IPR004046; GST_Cterm.
DR   Pfam; PF00043; GST_C; 1.
KW   Protein biosynthesis.
SQ   SEQUENCE   320 AA;  35349 MW;  F253726B63C12BAB CRC64;

  Query Match             88.5%;  Score 1464;  DB 1;  Length 320;
  Best Local Similarity   87.2%;  Pred. No. 1.6e-112;
  Matches  279;  Conservative   16;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              |||||||||||| |||||||||||||||||| ::  ||  |||||| |  ||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||| |||||||||:|||||||  || ||||||||||||||||
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              |||||||||||||||||||||||| :|||||||||||||:||||:|||||| :|| | :
```

```
Db        121 DIVINANPASPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              ||||||||||||||||||||:||||||||||||||||||||||||||  ||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              :|||:||||||||||||||||||:|||| ||||||||||||||||| || |    | |||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWLKSCENLAPFSTALQLLK 320
              ||::|||||||:|||:|||
Db        301 RWMRSCENLAPFNTALKLLK 320


RESULT 5
Q96CZ5
ID   Q96CZ5      PRELIMINARY;      PRT;   320 AA.
AC   Q96CZ5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   JTV1.
GN   Name=JTV1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RA   Strausberg R.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC013630; AAH13630.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   320 AA;  35335 MW;  19F14BF758612E08 CRC64;

  Query Match             88.2%;  Score 1460;  DB 2;  Length 320;
  Best Local Similarity   86.9%;  Pred. No. 3.4e-112;
  Matches  278;  Conservative   16;  Mismatches   26;  Indels    0;  Gaps    0;

Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              |||||||||||| |||||||||||||||| ::  ||  |||||| |  ||||||||||
Db          1 MPMYQVKPYHGGGAPLRVELPTCMYRLPNVHGRSYGPAPGAGHVQEESNLSLQALESRQD 60

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||| |||||||||||:|||||||| || ||||||||||||||||
```

```
Db         61 DILKRLYELKAAVDGLSKMIQTPDADLDVTNIIQADEPTTLTTNALDLNSVLGKDYGALK 120

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              |||||||| |||||||||||||| :||||||||||||:|||||:|||||| :|| | :
Db        121 DIVINANPGSPPLSLLVLHRLLCEHFRVLSTVHTHSSVKSVPENLLKCFGEQNKKQPRQD 180

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              ||||||||||||||||||||:||||||||||||||||||||||||||  ||||||||||
Db        181 YQLGFTLIWKNVPKTQMKFSIQTMCPIEGEGNIARFLFSLFGQKHNAVNATLIDSWVDIA 240

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              :|||:||||||||||||||||||:|||| ||||||||||||||||| || |    | |||
Db        241 IFQLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVVLWSVLQQIGGCSVTVPANVQ 300

Qy        301 RWLKSCENLAPFSTALQLLK 320
              ||::|||||||:|||:|||
Db        301 RWMRSCENLAPFNTALKLLK 320


RESULT 6
Q8R3V2
ID   Q8R3V2      PRELIMINARY;      PRT;   280 AA.
AC   Q8R3V2;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   JTV1.
GN   Name=Jtv1;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RA   Strausberg R.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC024480; AAH24480.1; -.
DR   MGD; MGI:2385237; Jtv1.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   280 AA;  31097 MW;  E5BE81498983FD2E CRC64;

  Query Match             85.3%;  Score 1412;  DB 2;  Length 280;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-108;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         45 QETSEPSLQALESRQDDILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATN 104
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 QETSEPSLQALESRQDDILKRLYELKAAVDGLSKMIHTPDADLLDVTNILQADEPTTLATN 64

Qy        105 TLDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPEN 164
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 TLDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPEN 124

Qy        165 LVKCFGEQARKQSRHEYQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQK 224
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 LVKCFGEQARKQSRHEYQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQK 184

Qy        225 HNAVTLTLIDSWVDIAMFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSV 284
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 HNAVTLTLIDSWVDIAMFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSV 244

Qy        285 LQQTGGSSGAAPTNVQRWLKSCENLAPFSTALQLLK 320
                 ||||||||||||||||||||||||||||||||||
Db        245 LQQTGGSSGAAPTNVQRWLKSCENLAPFSTALQLLK 280


RESULT 7
Q6DK86
ID   Q6DK86      PRELIMINARY;      PRT;   311 AA.
AC   Q6DK86;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   MGC69221 protein.
GN   Name=MGC69221;
OS   Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8364;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C ,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Gerhard D.S.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC074561; AAH74561.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   311 AA;  34480 MW;  D98F27F73C466154 CRC64;

  Query Match           69.7%;  Score 1153.5;  DB 2;  Length 311;
  Best Local Similarity   67.5%;  Pred. No. 6.8e-87;
  Matches  216;  Conservative   49;  Mismatches   46;  Indels    9;  Gaps    3;

Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              ||||:::|| ||   ::|:|||||||||||||   :       :::: ::|::|||||||:
Db          1 MPMYRMQPYCGGE--IQVDLPTCMYRLPNVHQAVSE------NLEQAADPAIQALESRQE 52

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||| |||||||||:|:| || :       ||:|||||:||||:
Db         53 DILKRLYELKAAVDGLSKMIQTPDADLDVTDIIQTDEQSAPGPVVADLDSVLGKEYGALR 112

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              ||||||||: |||||||:|| |||||||||| |||||: ::|| |:||||:| : : ||:
Db        113 DIVINANPSLPPLSLLILHSLLCERYQVLSAVHTHSSIASIPEPLMKCFGDQLKMKPRHQ 172

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              ||||||||||:||| |||||:| ||||||||||||||||| | | ||:  |||| ||| |
Db        173 YQLGFTLIWKDVPKPQMKFSIQNMCPIEGEGNIARFLFSLLGYKFNAINATLIDGWVDTA 232

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              :|||||||||||||| ::||:|||:|||||||||||||:| |  :|| | |: | | :||
Db        233 IFQLREGSSKEKAAVLKAMNTALGKSPWLVGNELTVADIVSWCAVQQCGNST-AVPPSVQ 291

Qy        301 RWLKSCENLAPFSTALQLLK 320
              :|:||||||| | : |: ||
Db        292 KWMKSCENLASFKSVLRFLK 311


RESULT 8
Q6INU4
ID   Q6INU4      PRELIMINARY;      PRT;   311 AA.
AC   Q6INU4;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   MGC80304 protein.
GN   Name=MGC80304;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
```

```
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RA   Klein S., Gerhard D.S.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC072178; AAH72178.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   311 AA;  34424 MW;  D5E8325C18D88751 CRC64;

  Query Match             69.3%;  Score 1146.5;  DB 2;  Length 311;
  Best Local Similarity   68.1%;  Pred. No. 2.6e-86;
  Matches  218;  Conservative   44;  Mismatches   49;  Indels    9;  Gaps    3;

Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              ||||:::|| ||   ::|:|||||||||||| |      | : :: ::|::|||||||:
Db          1 MPMYRMQPYCGGE--IQVDLPTCMYRLPNVHRAT------AENHEQAADPAIQALESRQE 52

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||| |||||||||:|:|||| :   :   ||: ||| |||||:
Db         53 DILKRLYELKAAVDGLSKMIQTPDADLDVTDIIQADEQSATGSVVADLDVVLGNDYGALR 112

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              ||||||||:  |||||:|| ||||||:||| |||||| :::|| |:||||:| : : |||
Db        113 DIVINANPSLAPLSLLILHSLLCERYQVLSAVHTHSSITSIPEPLLKCFGDQLKMKPRHE 172

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              ||||||||||:||| |||||:| |||||||||| |||||| |   |||  |||| ||| |
Db        173 YQLGFTLIWKDVPKPQMKFSIQNMCPIEGEGNIGRFLFSLLGYTFNAVNATLIDGWVDTA 232

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              :|||||||||||||| ::||:|||:|||||||||||||:| |  :|| | |: | | |||
Db        233 IFQLREGSSKEKAAVLKAMNTALGKSPWLVGNELTVADIVSWCAVQQCGNST-AVPPNVQ 291

Qy        301 RWLKSCENLAPFSTALQLLK 320
              :|:||||||| | : |: ||
Db        292 KWMKSCENLASFKSVLRFLK 311


RESULT 9
Q7ZYD7
ID   Q7ZYD7      PRELIMINARY;      PRT;   311 AA.
AC   Q7ZYD7;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Jtv1-prov protein.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC043832; AAH43832.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   311 AA;  34647 MW;  2F08C9DA60DD63BE CRC64;

  Query Match             69.3%;  Score 1146.5;  DB 2;  Length 311;
  Best Local Similarity   68.1%;  Pred. No. 2.6e-86;
  Matches  218;  Conservative   44;  Mismatches   49;  Indels    9;  Gaps    4;

Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRQD 60
              ||||:::|| ||   ::|:|||||||| |||      ||   | :: ::|::|||||||:
Db          1 MPMYRMQPYCGGE--IQVDLPTCMYRLANVHQ-----ATSENH-EQAADPAIQALESRQE 52

Qy         61 DILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALK 120
              |||||||||||||||||| |||||||||:|:| || :   :   ||: ||||:||||:
Db         53 DILKRLYELKAAVDGLSKMIQTPDADLDVTHIIQTDEHSAPGSVVADLDLVLGKEYGALR 112

Qy        121 DIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFGEQARKQSRHE 180
              ||||||||: ||||||:|| ||||||:||| |||||| :::|| |:|||||  : : |||
Db        113 DIVINANPSLPPLSLLILHSLLCERYQVLSAVHTHSSITSIPEPLLKCFGEQLKMKLRHE 172

Qy        181 YQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTLIDSWVDIA 240
              ||||||||||:||| |||||:| |||||||||| |||||||| | |||  |||| ||| |
Db        173 YQLGFTLIWKDVPKPQMKFSIQNMCPIEGEGNIGRFLFSLFGYKFNAVNATLIDGWVDTA 232

Qy        241 MFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSSGAAPTNVQ 300
              :||||:|||||||| ::||: ||:|||||||||||||:| |  :|| | |:   | |||
Db        233 IFQLRDGSSKEKAAVLKAMNTTLGKSPWLVGNELTVADIVSWCAVQQCGNST-VVPPNVQ 291

Qy        301 RWLKSCENLAPFSTALQLLK 320
              :|:||||||| | : |: ||
Db        292 KWMKSCENLASFKSVLRFLK 311


RESULT 10
Q7T3C0
ID   Q7T3C0      PRELIMINARY;      PRT;   321 AA.
AC   Q7T3C0;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Zgc:63976.
GN   ORFNames=zgc:63976;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
```

```
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Strausberg R.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC053178; AAH53178.1; -.
DR   ZFIN; ZDB-GENE-040426-2652; zgc:63976.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   321 AA;  34852 MW;  BA8F6B951208244A CRC64;

  Query Match             56.0%;  Score 926.5;  DB 2;  Length 321;
  Best Local Similarity   57.3%;  Pred. No. 4.1e-68;
  Matches  188;  Conservative   48;  Mismatches   77;  Indels   15;  Gaps    7;

Qy          1 MPMYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTS---PATDAGHVQETSEPSLQALES 57
              ||||||||     | : |:||||||:|||||:  |     |   | |    :|:::|||
Db          1 MPMYQVKPV--SPADITVDLPTCMYKLPNVHAQGASLGEHALQNGEV----DPTVKALEE 54

Qy         58 RQDDILKRLYELKAAVDGLSKMIHTPDADLDVTNI---LQADEPTILATNTLDLNSVLGK 114
              |||:||::|||||| ||||:| : ||||||| : :   |       :   | ||: :|||
Db         55 RQDEILRKLYELKATVDGLAKTVTTPDADLDASTLAHTLTHTPADAVLRGTADLDDLLGK 114

Qy        115 DYGALKDIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNVPENLVKCFG-EQA 173
              | |||:|||||||| |||||||| |||:|::|||:|| ||||  ||  |: | |
Db        115 DLGALRDIVINANPAQPPLSLLVLHALLCQRFQVLSSVHVHSSVSTVPAPLLSCLGPRHT 174

Qy        174 RKQSRHEYQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFG-QKHNAVTLTL 232
                :|| :|||||||||:| | ||||| | ||||||||||:||||: | | :  : |: ||
Db        175 HSYARHRFQLGFTLIWKDVSKLQMKFSTQNMCPIEGEGNVARFLYRLLGAEPRDPVSATL 234

Qy        233 IDSWVDIAMFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSVLQQTGGSS 292
              :| ||| |:||| || |||:||| |::|:||||||||:| | ::||:|    : ||| :|
Db        235 MDGWVDTALFQLAEGGSKERAAVLRALNAALGRSPWLLGQEFSLADIVSACCVLQTGQTS 294

Qy        293 GAAPTNVQRWLKSCENLAPFSTALQLLK 320
               :|| |||||||||:||  ||    ||:
Db        295 -SAPANVQRWLKSCQNLGYFSCVDPLLQ 321


RESULT 11
Q7Q7A3
ID   Q7Q7A3      PRELIMINARY;      PRT;   340 AA.
AC   Q7Q7A3;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   AgCP5808 (Fragment).
GN   Name=agCG50514; ORFNames=ENSANGG00000011827;
OS   Anopheles gambiae str. PEST.
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX   NCBI_TaxID=180454;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PEST;
RA   Anopheles Genome Sequencing Consortium;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AAAB01008960; EAA10870.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   340 AA;  37467 MW;  FD0D8525CD4EB976 CRC64;

  Query Match             14.9%;  Score 247;  DB 2;  Length 340;
  Best Local Similarity   28.7%;  Pred. No. 6.1e-12;
  Matches   94;  Conservative   49;  Mismatches  118;  Indels   66;  Gaps   13;

Qy          3 MYQVKPYHGGSAPLRVELPTCMYRLPNVHSKTTSP-------ATD---AGHV-------- 44
              ||::||    :  :  ||||||| |  | : :  |        ||    || |
Db         12 MYRLKPVM--NQDVVCELPTCMYTLKPVCAYSHDPKAFSADSATGLAAAGPVDPMSIEID 69

Qy         45 ---QETSEPSLQALESRQDDILKRLYELKAAVDGLSKMIHTPDADLDV-TNILQADEPTT 100
                 |:     |: |  ||  :|::| |||       | |     :| :  |   | :|:|
Db         70 NLLQKAENDELRMLAERQQRVLQQLAELK-------KEIMAMRTELKLNANAPPAVQPST 122

Qy        101 -------LATNTLDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCERYRVLSTVH 153
                    |    ::|          |:| |:||:|   | ||| |  |  :|  :
Db        123 PLKSKAQLKAEPINLT--------CLQDFVVNASPEYVPYSLLALKNLWKDRLNLQVECF 174

Qy        154 THSSVKNVPE------NLVKCFGEQARKQSRHEYQLGFTLIWKNV-PKTQMKFSVQTMCP 206
              |||:|  : |      | |   |  |    |    :  |||||||   |:|  |  :  |
Db        175 THSTVPKLSEEALAFQNAVTASGTAAANLPR----IKVTLIWKNVGAYTEMITSPTSYVP 230

Qy        207 IEGEGNIARFLFSL------FGQKHNAVTLTLIDSWVDIAMFQLREGSSKEKAAVFRSMN 260
              | || || |:|         : |: |   :  :|| :|      | : : |::  : |::
Db        231 ICGEVNILRYLSRCGPSEFNYEQQDN---VDEVDSILDACYLLLNKQNVKQRQQILRTLG 287

Qy        261 SALGRSPWLVGNELTVADVVLWSVLQQ 287
              : ||::    | :|:: |:   | ::|
Db        288 AKLGKAAGFGGADLSLCDIAFTSAVKQ 314


RESULT 12
Q7KUM5
ID   Q7KUM5      PRELIMINARY;      PRT;   301 AA.
AC   Q7KUM5;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   CG12304-PB.
GN   ORFNames=CG12304;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA   Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
```

```
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA   Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426065; PubMed=12537568;
RA   Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA   Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA   George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA   Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA   Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA   Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT   "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT   melanogaster euchromatic genome sequence.";
RL   Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426070; PubMed=12537573;
RA   Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA   Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA   Ashburner M., Celniker S.E.;
RT   "The transposable elements of the Drosophila melanogaster euchromatin:
RT   a genomics perspective.";
RL   Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfied E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN   [5]
RP   SEQUENCE FROM N.A.
RG   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RG   FlyBase;
RL   Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003530; AAS64998.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
SQ   SEQUENCE   301 AA;  33241 MW;  35224E17FE03F3BE CRC64;

  Query Match             13.5%;  Score 224;  DB 2;  Length 301;
  Best Local Similarity   28.6%;  Pred. No. 4.1e-10;
  Matches   86;  Conservative   47;  Mismatches  122;  Indels   46;  Gaps    9;

Qy          3 MYQVKPYHGGSAPLRVELPTCMYRLPNV--------HSKTTSPATDAGHVQETSEPSLQA 54
              ||::|          ::||||||  | ||           :|| :| |     :|     :
Db          1 MYELKTL---LPQFDIKLPTCMYPLKNVSLAADSLASGSSTSASTSA---STSSCDDTAS 54

Qy         55 LESRQDDILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGK 114
               : :||: :||:| |||| :  :         | | |     :  : ||    |       |
Db         55 VAARQEKVLKQLEELKAQLGQIR-------AGLGVCG--KTFQHTTAFQNG-------GL 98

Qy        115 DYGALKDIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNV-------PENLVK 167
                |:|:||| :|    | :|| |       | :     |||:: ::          || |
Db         99 KEVPLQDVVINGHPNFIPYALLALKNAWRNLYTIDVKTFTHSTMADIGPAAREFEANLAK 158

Qy        168 CFGEQARKQSRHEYQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNA 227
                  |        ::  ||||||   |:|  |     || || || |:|  :   ::
Db        159 VPVNPALP------KISVTLIWKNCEHTEMISSPTMYVPIYGEVNIIRYLGRVGPAEYRY 212

Qy        228 VTLTL---IDSWVDIAMFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSV 284
                |    ||  :||    ||  : | : |: | ::  | :  :  |::::|||| ::|
Db        213 EGSPLCNEIDLVLDICYQLLRCNTHKTQVAMVRLLDKRLQKQQYFGGSQMSVADVGVYSS 272

Qy        285 L 285
               |
Db        273 L 273


RESULT 13
Q8T060
ID   Q8T060      PRELIMINARY;      PRT;   313 AA.
AC   Q8T060;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   LD25772p (Fragment).
GN   ORFNames=CG12304;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA   George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA   Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA   Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA   Celniker S.;
RL   Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY069537; AAL39682.2; -.
DR   FlyBase; FBgn0036515; CG12304.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   313 AA;  34522 MW;  5AF9C801877F8531 CRC64;

  Query Match             13.5%;  Score 224;  DB 2;  Length 313;
  Best Local Similarity   28.6%;  Pred. No. 4.3e-10;
  Matches   86;  Conservative   47;  Mismatches  122;  Indels   46;  Gaps    9;

Qy          3 MYQVKPYHGGSAPLRVELPTCMYRLPNV--------HSKTTSPATDAGHVQETSEPSLQA 54
              ||::|          ::||||||  | ||           :|| :| |     :|     :
Db         13 MYELKTL---LPQFDIKLPTCMYPLKNVSLAADSLASGSSTSASTSA---STSSCDDTAS 66

Qy         55 LESRQDDILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGK 114
               : :||: :||:| |||| :  :         | | |     :  : ||    |       |
Db         67 VAARQEKVLKQLEELKAQLGQIR-------AGLGVCG--KTFQHTTAFQNG-------GL 110

Qy        115 DYGALKDIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNV-------PENLVK 167
                |:|:||| :|    | :|| |       | :     |||:: ::          || |
Db        111 KEVPLQDVVINGHPNFIPYALLALKNAWRNLYTIDVKTFTHSTMADIGPAAREFEANLAK 170
```

```
Qy        168 CFGEQARKQSRHEYQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNA 227
Db        171 VPVNPALP------KISVTLIWKNCEHTEMISSPTMYVPIYGEVNIIRYLGRVGPAEYRY 224

Qy        228 VTLTL---IDSWVDIAMFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSV 284
Db        225 EGSPLCNEIDLVLDICYQLLRCNTHKTQVAMVRLLDKRLQKQQYFGGSQMSVADVGVYSS 284

Qy        285 L 285
Db        285 L 285


RESULT 14
Q6NKM4
ID   Q6NKM4      PRELIMINARY;      PRT;   322 AA.
AC   Q6NKM4;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   LP12114p (Fragment).
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Carlson J., Chavez C., Frise E., George R., Pacleb J.,
RA   Park S., Wan K., Yu C., Rubin G.M., Celniker S.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BT012670; AAT08476.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR010987; GST_C_like.
DR   Pfam; PF00043; GST_C; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   322 AA;  35528 MW;  82A15927DF928DAE CRC64;

  Query Match           13.5%;  Score 224;  DB 2;  Length 322;
  Best Local Similarity   28.6%;  Pred. No. 4.5e-10;
  Matches   86;  Conservative   47;  Mismatches  122;  Indels   46;  Gaps    9;

Qy          3 MYQVKPYHGGSAPLRVELPTCMYRLPNV--------HSKTTSPATDAGHVQETSEPSLQA 54
Db         22 MYELKTL---LPQFDIKLPTCMYPLKNVSLAADSLASGSSTSASTSA---STSSCDDTAS 75

Qy         55 LESRQDDILKRLYELKAAVDGLSKMIHTPDADLDVTNILQADEPTTLATNTLDLNSVLGK 114
Db         76 VAARQEKVLKQLEELKAQLGQIR-------AGLGVCG--KTFQHTTAFQNG-------GL 119

Qy        115 DYGALKDIVINANPASPPLSLLVLHRLLCERYRVLSTVHTHSSVKNV-------PENLVK 167
Db        120 KEVPLQDVVINGHPNFIPYALLALKNAWRNLYTIDVKTFTHSTMADIGPAARERFEANLAK 179

Qy        168 CFGEQARKQSRHEYQLGFTLIWKNVPKTQMKFSVQTMCPIEGEGNIARFLFSLFGQKHNA 227
Db        180 VPVNPALP------KISVTLIWKNCEHTEMISSPTMYVPIYGEVNIIRYLGRVGPAEYRY 233

Qy        228 VTLTL---IDSWVDIAMFQLREGSSKEKAAVFRSMNSALGRSPWLVGNELTVADVVLWSV 284
Db        234 EGSPLCNEIDLVLDICYQLLRCNTHKTQVAMVRLLDKRLQKQQYFGGSQMSVADVGVYSS 293

Qy        285 L 285
Db        294 L 294


RESULT 15
MCA2_DROME
ID   MCA2_DROME     STANDARD;      PRT;   334 AA.
AC   Q9VUR3;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   25-JAN-2005 (Rel. 46, Last annotation update)
DE   Probable multisynthetase complex auxiliary component p38.
GN   ORFNames=CG12304;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   GENOME REANNOTATION.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
CC   -!- FUNCTION: Probable core protein of the multisynthetase complex
CC       that serves as a template for the assembly of the supramolecular
CC       structure (By similarity).
CC   -!- SUBUNIT: Component of the multisynthetase complex which is
CC       comprised of a bifunctional glutamyl-prolyl-tRNA synthetase, the
CC       monospecific isoleucyl, leucyl, glutaminyl, methionyl, lysyl,
CC       arginyl, and aspartyl-tRNA synthetases as well as three auxiliary
CC       proteins, p18, p48 and p43 (By similarity).
CC   -!- SIMILARITY: Contains 1 GST-like domain.
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE003530; AAF49612.1; -.
DR   IntAct; Q9VUR3; -.
DR   FlyBase; FBgn0036515; CG12304.
DR   InterPro; IPR010987; GST_C_like.
DR   InterPro; IPR004046; GST_Cterm.
DR   Pfam; PF00043; GST_C; 1.
KW   Protein biosynthesis.
FT   DOMAIN      280    327       GST-like.
SQ   SEQUENCE   334 AA;  36933 MW;  B68FD70AE621990F CRC64;

  Query Match             13.1%;  Score 216.5;  DB 1;  Length 334;
  Best Local Similarity   26.0%;  Pred. No. 2e-09;
  Matches   86;  Conservative   48;  Mismatches  124;  Indels   73;  Gaps    9;

Qy          3 MYQVKPYHGGSAPLRVELPTCMYRLPNV--------HSKTTSPATDAG------------ 42
              ||::|          ::|||||| | ||           :|| :| |
Db          1 MYELKTL---LPQFDIKLPTCMYPLKNVSLAADSLASGSSTSASTSASTSSCKLEANRID 57

Qy         43 -----------------HVQETSEPSLQALESRQDDILKRLYELKAAVDGLSKMIHTPD 84
                               :|   :    :: :||: :||:| |||| :  :
Db         58 RTGRNAATCALDLDSLGRQIQRLLKDDTASVAARQEKVLKQLEELKAQLGQIR------- 110

Qy         85 ADLDVTNILQADEPTTLATNTLDLNSVLGKDYGALKDIVINANPASPPLSLLVLHRLLCE 144
              | | |    :  : ||   |         |     |:|:||| :|   | :|| |
Db        111 AGLGVCG--KTFQHTTAFQNG-------GLKEVPLQDVVINGHPNFIPYALLALKNAWRN 161

Qy        145 RYRVLSTVHTHSSVKNV-------PENLVKCFGEQARKQSRHEYQLGFTLIWKNVPKTQM 197
               | :     |||:: ::         || |     |        ::  ||||||   |:|
Db        162 LYTIDVKTFTHSTMADIGPAAREFEANLAKVPVNPALP------KISVTLIWKNCEHTEM 215

Qy        198 KFSVQTMCPIEGEGNIARFLFSLFGQKHNAVTLTL---IDSWVDIAMFQLREGSSKEKAA 254
               |     || || || |:|  :   ::      |   ||  :||    ||  : | : |
Db        216 ISSPTMYVPIYGEVNIIRYLGRVGPAEYRYEGSPLCNEIDLVLDICYQLLRCNTHKTQVA 275

Qy        255 VFRSMNSALGRSPWLVGNELTVADVVLWSVL 285
              : | ::  | :  :  |::::|||| ::| |
Db        276 MVRLLDKRLQKQQYFGGSQMSVADVGVYSSL 306


Search completed: February 23, 2005, 13:57:15
Job time : 61.9811 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:25:39 ; Search time 102.258 Seconds
(without alignments)
1758.725 Million cell updates/sec

Title: US-10-622-817-8
Perfect score: 2596
Sequence: 1 MIVFVRFNSSHGFPVEVDSD.....CWNCGCEWNRVCMGDHWFDV 465

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2596	100.0	465	2	AAY32501	Aay32501 Human par
2	2596	100.0	465	6	AAE30800	Aae30800 Human Par
3	2596	100.0	465	6	ABO07156	Abo07156 Human p53
4	2584	99.5	465	6	AAE30801	Aae30801 Human Par
5	2405	92.6	437	2	AAY32502	Aay32502 Human par
6	2405	92.6	437	6	ABO07157	Abo07157 Human p53
7	2170.5	83.6	464	4	AAB67517	Aab67517 Amino aci
8	2165.5	83.4	464	4	AAB67531	Aab67531 Amino aci
9	2165.5	83.4	464	4	AAB67532	Aab67532 Amino aci
10	2087.5	80.4	451	4	AAB67533	Aab67533 Amino aci
11	1696.5	65.4	316	6	ABO07158	Abo07158 Human p53
12	1679.5	64.7	344	4	AAB67521	Aab67521 Amino aci
13	1234.5	47.6	296	4	AAB67526	Aab67526 Amino aci
14	1058.5	40.8	468	4	ABB65114	Abb65114 Drosophil
15	1002	38.6	250	4	AAB67519	Aab67519 Amino aci
16	997	38.4	262	4	AAB67518	Aab67518 Amino aci
17	870	33.5	156	4	AAG67213	Aag67213 Amino aci
18	671.5	25.9	183	4	AAB67525	Aab67525 Amino aci
19	671.5	25.9	194	4	AAB67524	Aab67524 Amino aci
20	544.5	21.0	386	4	AAB86951	Aab86951 C. elegan
21	506	19.5	153	4	AAB67523	Aab67523 Amino aci
22	335	12.9	77	4	AAB67529	Aab67529 Amino aci
23	263	10.1	46	4	AAM17055	Aam17055 Peptide #
24	263	10.1	46	4	ABB36052	Abb36052 Peptide #
25	263	10.1	46	4	AAM29548	Aam29548 Peptide #
26	263	10.1	46	4	ABB30873	Abb30873 Peptide #
27	263	10.1	46	4	ABB21452	Abb21452 Protein #
28	263	10.1	46	4	AAM69222	Aam69222 Human bon
29	263	10.1	46	4	AAM56839	Aam56839 Human bra
30	263	10.1	46	4	ABG50898	Abg50898 Human liv
31	263	10.1	46	5	ABG38830	Abg38830 Human pep
32	252	9.7	106	8	ADM96489	Adm96489 Modified
33	251	9.7	63	4	AAB67522	Aab67522 Amino aci
34	251	9.7	105	4	AAB67520	Aab67520 Amino aci
35	251	9.7	105	8	ADM96487	Adm96487 Modified
36	240.5	9.3	503	4	ABB61708	Abb61708 Drosophil
37	240.5	9.3	503	7	ADB67082	Adb67082 Ariadne-1
38	240	9.2	511	4	ABB63665	Abb63665 Drosophil
39	239	9.2	520	8	ADP22562	Adp22562 Sea-squir
40	236.5	9.1	445	3	AAB21034	Aab21034 Human nuc
41	236.5	9.1	445	8	ADQ96154	Adq96154 T cell ac
42	236.5	9.1	445	8	ADQ96026	Adq96026 T cell ac
43	236.5	9.1	557	3	AAY98059	Aay98059 Human Rin
44	236.5	9.1	557	8	ADQ96028	Adq96028 T cell ac
45	236	9.1	53	4	AAB67528	Aab67528 Amino aci

ALIGNMENTS

RESULT 1
AAY32501
ID AAY32501 standard; protein; 465 AA.
XX
AC AAY32501;
XX
DT 21-OCT-1999 (first entry)
XX
DE Human parkin gene variant protein.
XX
KW Parkinson's disease related gene; parkin gene; variant; gene therapy.
XX
OS Homo sapiens.
XX
PN WO9940191-A1.
XX
PD 12-AUG-1999.
XX
PF 09-FEB-1999; 99WO-JP000545.
XX
PR 09-FEB-1998; 98JP-00027531.
XX
PA (SHIM/) SHIMIZU N.
PA (MIZU/) MIZUNO Y.
XX
PI Shimizu N, Mizuno Y;
XX
DR WPI; 1999-494295/41.
DR N-PSDB; AAX99923.
XX
PT Gene implicated in the pathology of Parkinson's disease, used for
PT treatment of the disease.
XX
PS Claim 1; Page 83-88; 114pp; English.
XX
CC This sequence is encoded by a gene of the invention, and is implicated in
CC the pathology of Parkinson's disease. This sequence is a variant of the
CC parkin gene found in parkinson's disease patients. The sequences may be
CC used for the diagnosis, treatment (including gene therapy) and
CC investigation of Parkinson's disease
XX
SQ Sequence 465 AA;

Query Match 100.0%; Score 2596; DB 2; Length 465;
Best Local Similarity 100.0%; Pred. No. 2.3e-219;
Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

```
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
Db        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
Db        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
Db        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
Db        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
Db        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
Db        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
```

RESULT 2
AAE30800
ID AAE30800 standard; protein; 465 AA.
XX
AC AAE30800;
XX
DT 24-FEB-2003 (first entry)
XX
DE Human Parkin protein.
XX
KW Human; Parkin protein; neurological disorder; apoptosis; gene therapy;
KW ischaemic stroke; Parkinson's disease; Alzheimer's disease; nootropic;
KW transgenic; cerebroprotective; neuroprotective; neurotransplantation.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Cleavage-site 126
XX
PN WO200279459-A2.
XX
PD 10-OCT-2002.
XX
PF 02-APR-2002; 2002WO-DK000221.
XX
PR 29-MAR-2001; 2001DK-00000525.
PR 03-APR-2001; 2001US-0281286P.
XX
PA (NSGE-) NSGENE AS.
XX
PI Jensen PH;
XX
DR WPI; 2003-046812/04.
DR N-PSDB; AAD47679.
XX
PT New isolated nucleic acid sequence encoding a Parkin polypeptide, useful
PT for treating, preventing or diagnosing neurological disorders, e.g.
PT Parkinson's disease, Alzheimer's disease or ischemic stroke, and in
PT screening assays.
XX
PS Claim 10; Page 69; 71pp; English.
XX
CC The invention relates to Parkin protein and its corresponding nucleic
CC acid sequence. The nucleic acid sequence is useful for altering the
CC proteolytic processing of Parkin at its potential cleavage site at Asp
CC 126. The invention is used in manufacturing or testing a pharmaceutical
CC composition for treating and/or preventing a neurological disorder, e.g.
CC Alzheimer's disease or ischaemic stroke. It also used for detecting the
CC occurrence of proteolytic processing of Parkin at Asp 126 in a sample, in
CC monitoring a potential disposition for a neurodegenerative disease, and
CC for treating, preventing and/or diagnosing Parkinson's disease or other
CC neurodegenerative disorders. The viral vector is used for transforming
CC neuronal cells in vivo or ex vivo. The invention is useful for
CC neurotransplantation into the CNS of a mammal. It may be used in
CC screening assays to identify compounds that increase or decrease
CC apoptosis. It is also used in gene therapy. The present sequence is human
CC Parkin protein
XX
SQ Sequence 465 AA;

Query Match 100.0%; Score 2596; DB 6; Length 465;
Best Local Similarity 100.0%; Pred. No. 2.3e-219;
Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
Db        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
Db        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
Db        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
Db        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
Db        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
Db        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
```

RESULT 3
ABO07156
ID ABO07156 standard; protein; 465 AA.
XX
AC ABO07156;
XX
DT 13-AUG-2003 (first entry)
XX
DE Human p53 modifying protein, SEQ ID 116.
XX
KW Human; p53 modifier; cytostatic; cancer; cytostatic; antiangiogenic;
KW antiapoptotic; p53 pathway; breast cancer; colon cancer; kidney cancer;
KW lung cancer; ovarian cancer; angiogenesis; cell cycle;
KW apoptotic disorder; cell proliferation disorder.
XX
OS Homo sapiens.
XX

PN WO200299122-A1.
XX
PD 12-DEC-2002.
XX
PF 03-JUN-2002; 2002WO-US017382.
XX
PR 05-JUN-2001; 2001US-0296076P.
PR 10-OCT-2001; 2001US-0328605P.
PR 15-FEB-2002; 2002US-0357253P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX
DR WPI; 2003-156859/15.
DR N-PSDB; ACD13332.
XX
PT Identifying modulators of the p53 pathway for use in treating apoptotic
PT or cell proliferation disorders, comprises screening for agents that
PT modulate activity of a human ortholog of genes that modify the p53
PT pathway in Drosophila.
XX
PS Example 2; Page 399-401; 678pp; English.
XX
CC The invention relates to identifying (M1) a candidate p53 pathway
CC modulating agent, by contacting an assay system comprising a purified HM
CC polypeptide (human orthologue of genes that modify the p53 pathway in
CC Drosophila) or nucleic acid with a test agent under conditions, where but
CC for the presence of the test agent, the system provides a reference
CC activity, and detecting a test agent-biased activity of the assay system.
CC Also included are modulating (M2) a p53 pathway of a cell (comprising
CC contacting a cell defective in p53 function with a candidate modulator
CC that specifically binds to a HM polypeptide comprising an HM amino acid
CC sequence, where p53 function is restored), modulating (M3) a p53 pathway
CC in a mammalian cell (comprising contacting the cell with an agent that
CC specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)
CC a disease in a patient (comprising: (a) obtaining a biological sample
CC from the patient; (b) contacting the sample with a probe for HM
CC expression; (c) comparing the results with a control; and (d) determining
CC whether the comparison indicates a likelihood disease). (M1) is useful
CC for identifying modulators of the p53 pathway. A probe for HM expression
CC is useful for diagnosing breast, colon, kidney, lung and ovarian cancer,
CC in a patient, where the cancer has greater than 25 % expression level.
CC Modulators identified by (M1) are useful in a variety of diagnostic and
CC therapeutic applications, where disease or disorder prognosis is related
CC to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell
CC proliferation disorders (e.g. cancer). Another two new methods (M2 and
CC M3) are useful for modulating the p53 pathway of a cell, thus restoring
CC the p53 function of the cell, so that the cell undergoes normal
CC proliferation or progression through the cell cycle. (M2) and (M3) are
CC also useful for treating defects in the p53 pathway such as angiogenic,
CC apoptotic or cell proliferation disorders. The present sequence
CC represents a human p53 pathway modifying protein
XX
SQ Sequence 465 AA;

Query Match 100.0%; Score 2596; DB 6; Length 465;
Best Local Similarity 100.0%; Pred. No. 2.3e-219;
Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |||||||||||||||||||||||||||||||||||||||||||||
Db        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
```

RESULT 4
AAE30801
ID AAE30801 standard; protein; 465 AA.
XX
AC AAE30801;
XX
DT 24-FEB-2003 (first entry)
XX
DE Human Parkin D126 mutant protein.
XX
KW Human; Parkin protein; neurological disorder; apoptosis; gene therapy;
KW ischaemic stroke; Parkinson's disease; Alzheimer's disease; nootropic;
KW transgenic; cerebroprotective; neuroprotective; neurotransplantation;
KW mutant; mutein.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 126
FT /note= "Wild-type Asp is replaced with Glu"
FT Cleavage-site 126
FT Misc-difference 223
FT /note= "Encoded by CCA"
XX
PN WO200279459-A2.
XX
PD 10-OCT-2002.
XX
PF 02-APR-2002; 2002WO-DK000221.
XX
PR 29-MAR-2001; 2001DK-00000525.
PR 03-APR-2001; 2001US-0281286P.
XX
PA (NSGE-) NSGENE AS.
XX
PI Jensen PH;
XX
DR WPI; 2003-046812/04.
DR N-PSDB; AAD47680.
XX
PT New isolated nucleic acid sequence encoding a Parkin polypeptide, useful
PT for treating, preventing or diagnosing neurological disorders, e.g.
PT Parkinson's disease, Alzheimer's disease or ischemic stroke, and in
PT screening assays.
XX
PS Claim 10; Page 71; 71pp; English.
XX
CC The invention relates to Parkin protein and its corresponding nucleic
CC acid sequence. The nucleic acid sequence is useful for altering the
CC proteolytic processing of Parkin at its potential cleavage site at Asp
CC 126. The invention is used in manufacturing or testing a pharmaceutical
CC composition for treating and/or preventing a neurological disorder, e.g.

CC Alzheimer's disease or ischaemic stroke. It also used for detecting the
CC occurrence of proteolytic processing of Parkin at Asp 126 in a sample, in
CC monitoring a potential disposition for a neurodegenerative disease, and
CC for treating, preventing and/or diagnosing Parkinson's disease or other
CC neurodegenerative disorders. The viral vector is used for transforming
CC neuronal cells in vivo or ex vivo. The invention is useful for
CC neurotransplantation into the CNS of a mammal. It may be used in
CC screening assays to identify compounds that increase or decrease
CC apoptosis. It is also used in gene therapy. The present sequence is human
CC Parkin mutant protein
XX
SQ Sequence 465 AA;

Query Match 99.5%; Score 2584; DB 6; Length 465;
Best Local Similarity 99.6%; Pred. No. 2.6e-218;
Matches 463; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VILHTESRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETSVALHLIATNSRNITCIT 240

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |||||||||||||||||||||||||||||||||||||||||||||
Db        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
```

RESULT 5
AAY32502
ID AAY32502 standard; protein; 437 AA.
XX
AC AAY32502;
XX
DT 21-OCT-1999 (first entry)
XX
DE Human parkin gene variant protein.
XX
KW Parkinson's disease related gene; parkin gene; variant; gene therapy.
XX
OS Homo sapiens.
XX
PN WO9940191-A1.
XX
PD 12-AUG-1999.
XX
PF 09-FEB-1999; 99WO-JP000545.
XX
PR 09-FEB-1998; 98JP-00027531.
XX
PA (SHIM/) SHIMIZU N.
PA (MIZU/) MIZUNO Y.
XX
PI Shimizu N, Mizuno Y;
XX
DR WPI; 1999-494295/41.
DR N-PSDB; AAX99924.
XX
PT Gene implicated in the pathology of Parkinson's disease, used for
PT treatment of the disease.
XX
PS Claim 1; Page 89-94; 114pp; English.
XX
CC This sequence is encoded by a gene of the invention, and is implicated in
CC the pathology of Parkinson's disease. This sequence is a variant of the
CC parkin gene found in parkinson's disease patients. The sequences may be
CC used for the diagnosis, treatment (including gene therapy) and
CC investigation of Parkinson's disease
XX
SQ Sequence 437 AA;

Query Match 92.6%; Score 2405; DB 2; Length 437;
Best Local Similarity 94.0%; Pred. No. 1.3e-202;
Matches 437; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

```
Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQ-- 178

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 --------------------------EFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 212

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 272

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 332

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 392

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |||||||||||||||||||||||||||||||||||||||||||||
Db        393 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 437
```

RESULT 6
ABO07157
ID ABO07157 standard; protein; 437 AA.
XX
AC ABO07157;
XX
DT 13-AUG-2003 (first entry)
XX
DE Human p53 modifying protein, SEQ ID 117.
XX
KW Human; p53 modifier; cytostatic; cancer; cytostatic; antiangiogenic;
KW antiapoptotic; p53 pathway; breast cancer; colon cancer; kidney cancer;
KW lung cancer; ovarian cancer; angiogenesis; cell cycle;
KW apoptotic disorder; cell proliferation disorder.
XX
OS Homo sapiens.

XX
PN WO200299122-A1.
XX
PD 12-DEC-2002.
XX
PF 03-JUN-2002; 2002WO-US017382.
XX
PR 05-JUN-2001; 2001US-0296076P.
PR 10-OCT-2001; 2001US-0328605P.
PR 15-FEB-2002; 2002US-0357253P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX
DR WPI; 2003-156859/15.
DR N-PSDB; ACD13333.
XX
PT Identifying modulators of the p53 pathway for use in treating apoptotic
PT or cell proliferation disorders, comprises screening for agents that
PT modulate activity of a human ortholog of genes that modify the p53
PT pathway in Drosophila.
XX
PS Example 2; Page 401-402; 678pp; English.
XX
CC The invention relates to identifying (M1) a candidate p53 pathway
CC modulating agent, by contacting an assay system comprising a purified HM
CC polypeptide (human orthologue of genes that modify the p53 pathway in
CC Drosophila) or nucleic acid with a test agent under conditions, where but
CC for the presence of the test agent, the system provides a reference
CC activity, and detecting a test agent-biased activity of the assay system.
CC Also included are modulating (M2) a p53 pathway of a cell (comprising
CC contacting a cell defective in p53 function with a candidate modulator
CC that specifically binds to a HM polypeptide comprising an HM amino acid
CC sequence, where p53 function is restored), modulating (M3) a p53 pathway
CC in a mammalian cell (comprising contacting the cell with an agent that
CC specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)
CC a disease in a patient (comprising: (a) obtaining a biological sample
CC from the patient; (b) contacting the sample with a probe for HM
CC expression; (c) comparing the results with a control; and (d) determining
CC whether the comparison indicates a likelihood disease). (M1) is useful
CC for identifying modulators of the p53 pathway. A probe for HM expression
CC is useful for diagnosing breast, colon, kidney, lung and ovarian cancer,
CC in a patient, where the cancer has greater than 25 % expression level.
CC Modulators identified by (M1) are useful in a variety of diagnostic and
CC therapeutic applications, where disease or disorder prognosis is related
CC to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell
CC proliferation disorders (e.g. cancer). Another two new methods (M2 and
CC M3) are useful for modulating the p53 pathway of a cell, thus restoring
CC the p53 function of the cell, so that the cell undergoes normal
CC proliferation or progression through the cell cycle. (M2) and (M3) are
CC also useful for treating defects in the p53 pathway such as angiogenic,
CC apoptotic or cell proliferation disorders. The present sequence
CC represents a human p53 pathway modifying protein
XX
SQ Sequence 437 AA;

Query Match 92.6%; Score 2405; DB 6; Length 437;
Best Local Similarity 94.0%; Pred. No. 1.3e-202;
Matches 437; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

```
Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQ-- 178

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
                                           |||||||||||||||||||||||||||||||||
Db        179 ------------------------EFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 212

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 272

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 332

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 392

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |||||||||||||||||||||||||||||||||||||||||||||
Db        393 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 437
```

RESULT 7
AAB67517
ID AAB67517 standard; protein; 464 AA.
XX
AC AAB67517;
XX
DT 29-MAY-2001 (first entry)
XX
DE Amino acid sequence of a murine parkin2 polypeptide.
XX
KW Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW brain tumour; head trauma; stroke; vascular irregularity;
KW metabolic irregularity.
XX
OS Mus sp.
XX
PN EP1081225-A1.
XX
PD 07-MAR-2001.
XX
PF 30-AUG-1999; 99EP-00116766.
XX
PR 30-AUG-1999; 99EP-00116766.
XX
PA (BIOF-) BIOFRONTERA PHARM GMBH.
XX
PI Luebbert H;
XX
DR WPI; 2001-212797/22.
DR N-PSDB; AAF55244.
XX
PT New polynucleotides encoding mouse parkin2 protein, useful for producing
PT a transgenic non-human animal as an animal model for neurodegenerative
PT diseases.
XX
PS Disclosure; Page 17-19; 62pp; English.
XX
CC The present sequence represents a murine parkin2 polypeptide. Mutations
CC or deletions in the parkin2 gene cause Parkinson's disease in humans. The
CC human parkin2 gene is located in gene region 6q25.2-27. Parkin2
CC polypeptides and polynucleotides are useful for analysing
CC neurodegenerative diseases. They are also useful for testing the efficacy
CC of the treatment of a neurodegenerative disease such as Parkinson's
CC disease, Alzheimer's disease, Huntington's disease, amyotrophic lateral
CC sclerosis, Multi-system atrophy, Wilson's disease, Pick's disease, Prion
CC disease, and secondary causes inducing Parkinson's syndromes like toxins,
CC drugs, brain tumours, head trauma, stroke, vascular irregularities or
CC metabolic irregularities, associated with a less active or non-active
CC parkin protein
XX

SQ Sequence 464 AA;

Query Match 83.6%; Score 2170.5; DB 4; Length 464;
Best Local Similarity 83.4%; Pred. No. 6e-182;
Matches 388; Conservative 29; Mismatches 47; Indels 1; Gaps 1;

```
Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||:||||||||||| |||||||||||||||||||||||| |  ||||||
Db          1 MIVFVRFNSSYGFPVEVDSDTSILQLKEVVAKRQGVPADQLRVIFAGKELPNHLTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              |:||||||||||| |:  | ||:|||:|:: : |   | :|||||||||  || ||||||
Db         61 LEQQSIVHIVQRPRRRSHETNASGGDEPQSTSEGSIWESRSLTRVDLSSHTLPVDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||| |||::||   |  |  :  ||||::|||||| :|||||||| ||:|||||| |||
Db        121 VILDTDSKRDSEAARGPV-KPTYNSFFIYCKGPCHKVQPGKLRVQCGTCKQATLTLAQGP 179

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              |||||||||||||||||| |||| |||||||||||||:| |||:|| :| |:| ||
Db        180 SCWDDVLIPNRMSGECQSPDCPGTRAEFFFKCGAHPTSDKDTSVALNLITSNRRSIPCIA 239

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              ||||||||||||| |||||||||||||||||||||||||| ||||||||||||||||||
Db        240 CTDVRSPVLVFQCNHRHVICLDCFHLYCVTRLNDRQFVHDAQLGYSLPCVAGCPNSLIKE 299

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              |||||||||||| |||||||||||||||||||||||||||||||  ||||||||||||||
Db        300 LHHFRILGEEQYTRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEQGQRKVTCEGGNGLGC 359

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              || |||:|||||||:| :: | || |:|||||:||||||| ||||||||||||||||
Db        360 GFVFCRDCKEAYHEGDCDSLLEPSGATSQAYRVDKRAAEQARWEEASKETIKKTTKPCPR 419

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |:||:|||||||||||||:|||||||||||| |||||||||
Db        420 CNVPIEKNGGCMHMKCPQPQCKLEWCWNCGCEWNRACMGDHWFDV 464
```

RESULT 8
AAB67531
ID AAB67531 standard; protein; 464 AA.
XX
AC AAB67531;
XX
DT 29-MAY-2001 (first entry)
XX
DE Amino acid sequence of a mutated murine parkin2 polypeptide.
XX
KW Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW brain tumour; head trauma; stroke; vascular irregularity;
KW metabolic irregularity.
XX
OS Mus sp.
XX
PN EP1081225-A1.
XX
PD 07-MAR-2001.
XX
PF 30-AUG-1999; 99EP-00116766.
XX
PR 30-AUG-1999; 99EP-00116766.
XX
PA (BIOF-) BIOFRONTERA PHARM GMBH.
XX
PI Luebbert H;
XX
DR WPI; 2001-212797/22.
DR N-PSDB; AAF55258.
XX
PT New polynucleotides encoding mouse parkin2 protein, useful for producing
PT a transgenic non-human animal as an animal model for neurodegenerative
PT diseases.
XX
PS Claim 7; Page 47-49; 62pp; English.
XX
CC The present sequence represents a murine parkin2 polypeptide. The
CC sequence contains the mutation Lys161Asn. Mutations or deletions in the
CC parkin2 gene cause Parkinson's disease in humans. The human parkin2 gene
CC is located in gene region 6q25.2-27. Parkin2 polypeptides and
CC polynucleotides are useful for analysing neurodegenerative diseases. They
CC are also useful for testing the efficacy of the treatment of a
CC neurodegenerative disease such as Parkinson's disease, Alzheimer's
CC disease, Huntington's disease, amyotrophic lateral sclerosis, Multi-
CC system atrophy, Wilson's disease, Pick's disease, Prion disease, and
CC secondary causes inducing Parkinson's syndromes like toxins, drugs, brain
CC tumours, head trauma, stroke, vascular irregularities or metabolic
CC irregularities, associated with a less active or non-active parkin
CC protein
XX
SQ Sequence 464 AA;

Query Match 83.4%; Score 2165.5; DB 4; Length 464;
Best Local Similarity 83.2%; Pred. No. 1.7e-181;
Matches 387; Conservative 29; Mismatches 48; Indels 1; Gaps 1;

```
Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||:||||||||||| |||||||||||||||||||||||| |  ||||||
Db          1 MIVFVRFNSSYGFPVEVDSDTSILQLKEVVAKRQGVPADQLRVIFAGKELPNHLTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              |:||||||||||| |:  | ||:|||:|:: : |   | :|||||||||  || ||||||
Db         61 LEQQSIVHIVQRPRRRSHETNASGGDEPQSTSEGSIWESRSLTRVDLSSHTLPVDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||| |||::||   |  |  :  ||||::|||||| :|||| ||||| ||:|||||| |||
Db        121 VILDTDSKRDSEAARGPV-KPTYNSFFIYCKGPCHKVQPGNLRVQCGTCKQATLTLAQGP 179

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              |||||||||||||||||| |||| |||||||||||||:| |||:|| :| |:| ||
Db        180 SCWDDVLIPNRMSGECQSPDCPGTRAEFFFKCGAHPTSDKDTSVALNLITSNRRSIPCIA 239

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              ||||||||||||| |||||||||||||||||||||||||| ||||||||||||||||||
Db        240 CTDVRSPVLVFQCNHRHVICLDCFHLYCVTRLNDRQFVHDAQLGYSLPCVAGCPNSLIKE 299

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              |||||||||||| |||||||||||||||||||||||||||||||  ||||||||||||||
Db        300 LHHFRILGEEQYTRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEQGQRKVTCEGGNGLGC 359

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              || |||:|||||||:| :: | || |:|||||:||||||| ||||||||||||||||
Db        360 GFVFCRDCKEAYHEGDCDSLLEPSGATSQAYRVDKRAAEQARWEEASKETIKKTTKPCPR 419

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |:||:|||||||||||||:|||||||||||| |||||||||
Db        420 CNVPIEKNGGCMHMKCPQPQCKLEWCWNCGCEWNRACMGDHWFDV 464
```

RESULT 9
AAB67532
ID AAB67532 standard; protein; 464 AA.
XX
AC AAB67532;
XX
DT 29-MAY-2001 (first entry)
XX
DE Amino acid sequence of a mutated murine parkin2 polypeptide.
XX
KW Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW brain tumour; head trauma; stroke; vascular irregularity;
KW metabolic irregularity.

```
XX
OS   Mus sp.
XX
PN   EP1081225-A1.
XX
PD   07-MAR-2001.
XX
PF   30-AUG-1999;   99EP-00116766.
XX
PR   30-AUG-1999;   99EP-00116766.
XX
PA   (BIOF-) BIOFRONTERA PHARM GMBH.
XX
PI   Luebbert H;
XX
DR   WPI; 2001-212797/22.
DR   N-PSDB; AAF55259.
XX
PT   New polynucleotides encoding mouse parkin2 protein, useful for producing
PT   a transgenic non-human animal as an animal model for neurodegenerative
PT   diseases.
XX
PS   Claim 7; Page 49-51; 62pp; English.
XX
CC   The present sequence represents a murine parkin2 polypeptide. The
CC   sequence contains the mutation Thr415Asn. Mutations or deletions in the
CC   parkin2 gene cause Parkinson's disease in humans. The human parkin2 gene
CC   is located in gene region 6q25.2-27. Parkin2 polypeptides and
CC   polynucleotides are useful for analysing neurodegenerative diseases. They
CC   are also useful for testing the efficacy of the treatment of a
CC   neurodegenerative disease such as Parkinson's disease, Alzheimer's
CC   disease, Huntington's disease, amyotrophic lateral sclerosis, Multi-
CC   system atrophy, Wilson's disease, Pick's disease, Prion disease, and
CC   secondary causes inducing Parkinson's syndromes like toxins, drugs, brain
CC   tumours, head trauma, stroke, vascular irregularities or metabolic
CC   irregularities, associated with a less active or non-active parkin
CC   protein
XX
SQ   Sequence 464 AA;

  Query Match             83.4%;  Score 2165.5;  DB 4;  Length 464;
  Best Local Similarity   83.2%;  Pred. No. 1.7e-181;
  Matches  387;  Conservative   29;  Mismatches   48;  Indels    1;  Gaps    1;

Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              |||||||||||:|||||||||||||| |||||||||||||||||||||||| |  ||||||
Db          1 MIVFVRFNSSYGFPVEVDSDTSILQLKEVVAKRQGVPADQLRVIFAGKELPNHLTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              |:||||||||||| |:  | ||:|||:|:: : |    | :|||||||||  || ||||||
Db         61 LEQQSIVHIVQRPRRRSHETNASGGDEPQSTSEGSIWESRSLTRVDLSSHTLPVDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||| |||::||  |  |  :  ||||::|||||| :|||||||||| ||:|||||| |||
Db        121 VILDTDSKRDSEAARGPV-KPTYNSFFIYCKGPCHKVQPGKLRVQCGTCKQATLTLAQGP 179

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              |||||||||||||||||| |||| |||||||||||||:| |||:|| :| |:| ||
Db        180 SCWDDVLIPNRMSGECQSPDCPGTRAEFFFKCGAHPTSDKDTSVALNLITSNRRSIPCIA 239

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              |||||||||||||| ||||||||||||||||||||||||| ||||'|||||||||||||
Db        240 CTDVRSPVLVFQCNHRHVICLDCFHLYCVTRLNDRQFVHDAQLGYSLPCVAGCPNSLIKE 299

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              |||||||||||| |||||||||||||||||||||||||||||||  ||||||||||||||
Db        300 LHHFRILGEEQYTRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEQGQRKVTCEGGNGLGC 359

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              || |||:|||||||:| :: | || |:||||||:||||||| |||||||| |||||
Db        360 GFVFCRDCKEAYHEGDCDSLLEPSGATSQAYRVDKRAAEQARWEEASKETIKKTNKPCPR 419

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
```

```
              |:||:||||||||||||||:|||||||||||| |||||||||
Db        420 CNVPIEKNGGCMHMKCPQPQCKLEWCWNCGCEWNRACMGDHWFDV 464


RESULT 10
AAB67533
ID   AAB67533 standard; protein; 451 AA.
XX
AC   AAB67533;
XX
DT   29-MAY-2001  (first entry)
XX
DE   Amino acid sequence of a mutated murine parkin2 polypeptide.
XX
KW   Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW   Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW   Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW   brain tumour; head trauma; stroke; vascular irregularity;
KW   metabolic irregularity.
XX
OS   Mus sp.
XX
PN   EP1081225-A1.
XX
PD   07-MAR-2001.
XX
PF   30-AUG-1999;   99EP-00116766.
XX
PR   30-AUG-1999;   99EP-00116766.
XX
PA   (BIOF-) BIOFRONTERA PHARM GMBH.
XX
PI   Luebbert H;
XX
DR   WPI; 2001-212797/22.
DR   N-PSDB; AAF55260.
XX
PT   New polynucleotides encoding mouse parkin2 protein, useful for producing
PT   a transgenic non-human animal as an animal model for neurodegenerative
PT   diseases.
XX
PS   Claim 7; Page 51-53; 62pp; English.
XX
CC   The present sequence represents a murine parkin2 polypeptide. The
CC   sequence contains the mutation Trp453Stop. Mutations or deletions in the
CC   parkin2 gene cause Parkinson's disease in humans. The human parkin2 gene
CC   is located in gene region 6q25.2-27. Parkin2 polypeptides and
CC   polynucleotides are useful for analysing neurodegenerative diseases. They
CC   are also useful for testing the efficacy of the treatment of a
CC   neurodegenerative disease such as Parkinson's disease, Alzheimer's
CC   disease, Huntington's disease, amyotrophic lateral sclerosis, Multi-
CC   system atrophy, Wilson's disease, Pick's disease, Prion disease, and
CC   secondary causes inducing Parkinson's syndromes like toxins, drugs, brain
CC   tumours, head trauma, stroke, vascular irregularities or metabolic
CC   irregularities, associated with a less active or non-active parkin
CC   protein
XX
SQ   Sequence 451 AA;

  Query Match             80.4%;  Score 2087.5;  DB 4;  Length 451;
  Best Local Similarity   83.2%;  Pred. No. 1.2e-174;
  Matches  376;  Conservative   29;  Mismatches   46;  Indels    1;  Gaps    1;

Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              |||||||||||:|||||||||||||| |||||||||||||||||||||||| |  ||||||
Db          1 MIVFVRFNSSYGFPVEVDSDTSILQLKEVVAKRQGVPADQLRVIFAGKELPNHLTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              |:||||||||||| |:  | ||:|||:|:: : |    | :|||||||||  || ||||||
Db         61 LEQQSIVHIVQRPRRRSHETNASGGDEPQSTSEGSIWESRSLTRVDLSSHTLPVDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||| |||::||  |  |  :  ||||::|||||| :|||||||||| ||:|||||| |||
```

```
Db        121 VILDTDSKRDSEAARGPV-KPTYNSFFIYCKGPCHKVQPGKLRVQCGTCKQATLTLAQGP 179

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              |||||||||||||||||| |||| ||||||||||||||:| |||:|| :| |:| ||
Db        180 SCWDDVLIPNRMSGECQSPDCPGTRAEFFFKCGAHPTSDKDTSVALNLITSNRRSIPCIA 239

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              |||||||||||||| |||||||||||||||||||||||||| ||||||||||||||||||
Db        240 CTDVRSPVLVFQCNHRHVICLDCFHLYCVTRLNDRQFVHDAQLGYSLPCVAGCPNSLIKE 299

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              |||||||||||| |||||||||||||||||||||||||||||||  ||||||||||||||
Db        300 LHHFRILGEEQYTRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEQGQRKVTCEGGNGLGC 359

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              || |||:|||||||:| :: | || |:|||||:|||||||| ||||||||||||||||
Db        360 GFVFCRDCKEAYHEGDCDSLLEPSGATSQAYRVDKRAAEQARWEEASKETIKKTTKPCPR 419

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCE 452
              |:||:|||||||||||||||:||||||||||
Db        420 CNVPIEKNGGCMHMKCPQPQCKLEWCWNCGCE 451
```

RESULT 11
ABO07158
ID ABO07158 standard; protein; 316 AA.
XX
AC ABO07158;
XX
DT 13-AUG-2003 (first entry)
XX
DE Human p53 modifying protein, SEQ ID 118.
XX
KW Human; p53 modifier; cytostatic; cancer; cytostatic; antiangiogenic;
KW antiapoptotic; p53 pathway; breast cancer; colon cancer; kidney cancer;
KW lung cancer; ovarian cancer; angiogenesis; cell cycle;
KW apoptotic disorder; cell proliferation disorder.
XX
OS Homo sapiens.
XX
PN WO200299122-A1.
XX
PD 12-DEC-2002.
XX
PF 03-JUN-2002; 2002WO-US017382.
XX
PR 05-JUN-2001; 2001US-0296076P.
PR 10-OCT-2001; 2001US-0328605P.
PR 15-FEB-2002; 2002US-0357253P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX
DR WPI; 2003-156859/15.
DR N-PSDB; ACD13334.
XX
PT Identifying modulators of the p53 pathway for use in treating apoptotic
PT or cell proliferation disorders, comprises screening for agents that
PT modulate activity of a human ortholog of genes that modify the p53
PT pathway in Drosophila.
XX
PS Example 2; Page 402-403; 678pp; English.
XX
CC The invention relates to identifying (M1) a candidate p53 pathway
CC modulating agent, by contacting an assay system comprising a purified HM
CC polypeptide (human orthologue of genes that modify the p53 pathway in
CC Drosophila) or nucleic acid with a test agent under conditions, where but
CC for the presence of the test agent, the system provides a reference
CC activity, and detecting a test agent-biased activity of the assay system.
CC Also included are modulating (M2) a p53 pathway of a cell (comprising
CC contacting a cell defective in p53 function with a candidate modulator
CC that specifically binds to a HM polypeptide comprising an HM amino acid
CC sequence, where p53 function is restored), modulating (M3) a p53 pathway
CC in a mammalian cell (comprising contacting the cell with an agent that
CC specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)
CC a disease in a patient (comprising: (a) obtaining a biological sample
CC from the patient; (b) contacting the sample with a probe for HM
CC expression; (c) comparing the results with a control; and (d) determining
CC whether the comparison indicates a likelihood disease). (M1) is useful
CC for identifying modulators of the p53 pathway. A probe for HM expression
CC is useful for diagnosing breast, colon, kidney, lung and ovarian cancer,
CC in a patient, where the cancer has greater than 25 % expression level.
CC Modulators identified by (M1) are useful in a variety of diagnostic and
CC therapeutic applications, where disease or disorder prognosis is related
CC to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell
CC proliferation disorders (e.g. cancer). Another two new methods (M2 and
CC M3) are useful for modulating the p53 pathway of a cell, thus restoring
CC the p53 function of the cell, so that the cell undergoes normal
CC proliferation or progression through the cell cycle. (M2) and (M3) are
CC also useful for treating defects in the p53 pathway such as angiogenic,
CC apoptotic or cell proliferation disorders. The present sequence
CC represents a human p53 pathway modifying protein
XX
SQ Sequence 316 AA;

```
  Query Match             65.4%;  Score 1696.5;  DB 6;  Length 316;
  Best Local Similarity   68.0%;  Pred. No. 1.8e-140;
  Matches  316;  Conservative    0;  Mismatches    0;  Indels  149;  Gaps    1;

Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQ--- 57

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Db         58 ------------------------------------------------------------ 57

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180

Db         58 ------------------------------------------------------------ 57

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
                                              ||||||||||||||||||||||||||||||||||
Db         58 -------------------------EFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 91

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         92 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 151

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 211

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 271

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |||||||||||||||||||||||||||||||||||||||||||||
Db        272 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 316
```

RESULT 12
AAB67521
ID AAB67521 standard; protein; 344 AA.
XX
AC AAB67521;
XX
DT 29-MAY-2001 (first entry)
XX
DE Amino acid sequence of a murine truncated parkin2 polypeptide.
XX
KW Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;

```
KW   brain tumour; head trauma; stroke; vascular irregularity;
KW   metabolic irregularity.
XX
OS   Mus sp.
XX
PN   EP1081225-A1.
XX
PD   07-MAR-2001.
XX
PF   30-AUG-1999;   99EP-00116766.
XX
PR   30-AUG-1999;   99EP-00116766.
XX
PA   (BIOF-) BIOFRONTERA PHARM GMBH.
XX
PI   Luebbert H;
XX
DR   WPI; 2001-212797/22.
DR   N-PSDB; AAF55248.
XX
PT   New polynucleotides encoding mouse parkin2 protein, useful for producing
PT   a transgenic non-human animal as an animal model for neurodegenerative
PT   diseases.
XX
PS   Claim 7; Page 40-41; 62pp; English.
XX
CC   The present sequence represents a murine parkin2 polypeptide. The
CC   polynucleotide sequence contains a deletion, leading to a truncated
CC   protein. Mutations or deletions in the parkin2 gene cause Parkinson's
CC   disease in humans. The human parkin2 gene is located in gene region
CC   6q25.2-27. Parkin2 polypeptides and polynucleotides are useful for
CC   analysing neurodegenerative diseases. They are also useful for testing
CC   the efficacy of the treatment of a neurodegenerative disease such as
CC   Parkinson's disease, Alzheimer's disease, Huntington's disease,
CC   amyotrophic lateral sclerosis, Multi-system atrophy, Wilson's disease,
CC   Pick's disease, Prion disease, and secondary causes inducing Parkinson's
CC   syndromes like toxins, drugs, brain tumours, head trauma, stroke,
CC   vascular irregularities or metabolic irregularities, associated with a
CC   less active or non-active parkin protein
XX
SQ   Sequence 344 AA;

  Query Match             64.7%;  Score 1679.5;  DB 4;  Length 344;
  Best Local Similarity   66.0%;  Pred. No. 6.3e-139;
  Matches  307;  Conservative   14;  Mismatches   23;  Indels  121;  Gaps    1;

Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||:|||||||||||| |||||||||||||||||||||||| |  ||
Db          1 MIVFVRFNSSYGFPVEVDSDTSILQLKEVVAKRQGVPADQLRVIFAGKELPNHLTV---- 56

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Db         57 ------------------------------------------------------------ 56

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
                                                                  |||
Db         57 ------------------------------------------------------------QGP 59

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              |||||||||||||||||| |||| ||||||||||||||:| |||:|| :| |:| ||
Db         60 SCWDDVLIPNRMSGECQSPDCPGTRAEFFFKCGAHPTSDKDTSVALNLITSNRRSIPCIA 119

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              |||||||||||||| |||||||||||||||||||||||| ||||||||||||||||||
Db        120 CTDVRSPVLVFQCNHRHVICLDCFHLYCVTRLNDRQFVHDAQLGYSLPCVAGCPNSLIKE 179

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              |||||||||||| ||||||||||||||||||||||||||||  ||||||||||||||
Db        180 LHHFRILGEEQYTRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEQGQRKVTCEGGNGLGC 239

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              || |||:|||||||:| :: | || |:||||||:||||||||| ||||||||||||||||
Db        240 GFVFCRDCKEAYHEGDCDSLLEPSGATSQAYRVDKRAAEQARWEEASKETIKKTTKPCPR 299

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |:||:|||||||||||||||:|||||||||||||| |||||||||
Db        300 CNVPIEKNGGCMHMKCPQPQCKLEWCWNCGCEWNRACMGDHWFDV 344


RESULT 13
AAB67526
ID   AAB67526 standard; protein; 296 AA.
XX
AC   AAB67526;
XX
DT   29-MAY-2001  (first entry)
XX
DE   Amino acid sequence of a murine truncated parkin2 polypeptide.
XX
KW   Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW   Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW   Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW   brain tumour; head trauma; stroke; vascular irregularity;
KW   metabolic irregularity.
XX
OS   Mus sp.
XX
PN   EP1081225-A1.
XX
PD   07-MAR-2001.
XX
PF   30-AUG-1999;   99EP-00116766.
XX
PR   30-AUG-1999;   99EP-00116766.
XX
PA   (BIOF-) BIOFRONTERA PHARM GMBH.
XX
PI   Luebbert H;
XX
DR   WPI; 2001-212797/22.
DR   N-PSDB; AAF55253.
XX
PT   New polynucleotides encoding mouse parkin2 protein, useful for producing
PT   a transgenic non-human animal as an animal model for neurodegenerative
PT   diseases.
XX
PS   Claim 7; Page 44-45; 62pp; English.
XX
CC   The present sequence represents a murine parkin2 polypeptide. The
CC   polynucleotide sequence contains a deletion, leading to a truncated
CC   protein. Mutations or deletions in the parkin2 gene cause Parkinson's
CC   disease in humans. The human parkin2 gene is located in gene region
CC   6q25.2-27. Parkin2 polypeptides and polynucleotides are useful for
CC   analysing neurodegenerative diseases. They are also useful for testing
CC   the efficacy of the treatment of a neurodegenerative disease such as
CC   Parkinson's disease, Alzheimer's disease, Huntington's disease,
CC   amyotrophic lateral sclerosis, Multi-system atrophy, Wilson's disease,
CC   Pick's disease, Prion disease, and secondary causes inducing Parkinson's
CC   syndromes like toxins, drugs, brain tumours, head trauma, stroke,
CC   vascular irregularities or metabolic irregularities, associated with a
CC   less active or non-active parkin protein
XX
SQ   Sequence 296 AA;

  Query Match             47.6%;  Score 1234.5;  DB 4;  Length 296;
  Best Local Similarity   80.0%;  Pred. No. 7.3e-100;
  Matches  232;  Conservative   20;  Mismatches   37;  Indels    1;  Gaps    1;

Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||:|||||||||||| |||||||||||||||||||||||| |  ||||||
Db          1 MIVFVRFNSSYGFPVEVDSDTSILQLKEVVAKRQGVPADQLRVIFAGKELPNHLTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              |:||||||||||| |:  | ||:|||:|:: : |   | :|||||||||  || ||||||
Db         61 LEQQSIVHIVQRPRRRSHETNASGGDEPQSTSEGSIWESRSLTRVDLSSHTLPVDSVGLA 120
```

```
Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||| |||::||  |  |  :  ||||::|||||| :|||||||||| ||:|||||| |||
Db        121 VILDTDSKRDSEAARGPV-KPTYNSFFIYCKGPCHKVQPGKLRVQCGTCKQATLTLAQGP 179

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              |||||||||||||||||| |||| |||||||||||||:| |||:|| :| |:| ||
Db        180 SCWDDVLIPNRMSGECQSPDCPGTRAEFFFKCGAHPTSDKDTSVALNLITSNRRSIPCIA 239

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCV 290
              |||||||||||||| |||||||||||||||||||||||| |||||||||
Db        240 CTDVRSPVLVFQCNHRHVICLDCFHLYCVTRLNDRQFVHDAQLGYSLPCV 289
```

RESULT 14
ABB65114
ID ABB65114 standard; protein; 468 AA.
XX
AC ABB65114;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 22134.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL09217.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 22134; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 468 AA;

```
  Query Match             40.8%;  Score 1058.5;  DB 4;  Length 468;
  Best Local Similarity   42.6%;  Pred. No. 4e-84;
  Matches  201;  Conservative   80;  Mismatches  158;  Indels   33;  Gaps    9;

Qy          3 VFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCDLD 62
              ::|: |:     | ::    |  :||:|| : |:  | |::|||||| :  |:: |||
Db         18 IYVKTNTGKTLTVNLEPQWDIKNVKELVAPQLGLQPDDLKIIFAGKELSDATTIEQCDLG 77

Qy         63 QQSIVHIVQ-RPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLAV 121
                 |||::| :: ||  : |:: :   ::   :      :| : | :||
```

```
Db         78 QQSVLHAIRLRPPVQRQKIQSATLEEEEPSLSDEASKPLNETLLDLQ------------ 124

Qy        122 ILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGPS 181
              | :: | :          ::    |:|:|   | ::  |||||:|: |:    |: : |
Db        125 -LESEERLNITDEERVRAKA---HFFVHC-SQCDKLCNGKLRVRCALCKGGAFTVHRDPE 179

Qy        182 CWDDVLIPNRMSGECQSPH--CPGTS------AEFFFKCGAHPT-SDKETPVALHLIATN 232
              ||||||   |: | |:|    |   :      |||||||  | :  :|:    |:|| |
Db        180 CWDDVLKSRRIPGHCESLEVACVDNAAGDPPFAEFFFKCAEHVSGGEKDFAAPLNLIKNN 239

Qy        233 SRNITCITCTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAG 292
              :|: |: ||||   |||| | |:|| |:|||  || :|| :|||:  |  ||:||| ||
Db        240 IKNVPCLACTDVSDTVLVFPCASQHVTCIDCFRHYCRSRLGERQFMPHPDFGYTLPCPAG 299

Qy        293 CPNSLIKELHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTC 352
              | :| |:|:|||::|  |:|:|||::  || ||| ||||||:|||| ||| ||| |||||
Db        300 CEHSFIEEIHHFKLLTREEYDRYQRFATEEYVLQAGGVLCPQPGCGMGLLVEPDCRKVTC 359

Qy        353 EGGNGLGCGFAFCRECKEAYHEGEC-SAVFEASGTTTQAYRVDERAAEQARWEAASKETI 411
              :    |||: ||| | : || |||     || | :  | ||   | :|||: ||  ||
Db        360 QN----GCGYVFCRNCLQGYHIGECLPEGTGASATNSCEYTVDPNRAAEARWDEASNVTI 415

Qy        412 KKTTKPCPRCHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWF 463
              | :|||||:|  | |::|||||| | :  |  |||| |  || | ||| |||
Db        416 KVSTKPCPKCRTPTERDGGCMHMVCTRAGCGFEWCWVCQTEWTRDCMGAHWF 467
```

RESULT 15
AAB67519
ID AAB67519 standard; protein; 250 AA.
XX
AC AAB67519;
XX
DT 29-MAY-2001 (first entry)
XX
DE Amino acid sequence of a murine parkin2 polypeptide.
XX
KW Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW brain tumour; head trauma; stroke; vascular irregularity;
KW metabolic irregularity.
XX
OS Mus sp.
XX
PN EP1081225-A1.
XX
PD 07-MAR-2001.
XX
PF 30-AUG-1999; 99EP-00116766.
XX
PR 30-AUG-1999; 99EP-00116766.
XX
PA (BIOF-) BIOFRONTERA PHARM GMBH.
XX
PI Luebbert H;
XX
DR WPI; 2001-212797/22.
DR N-PSDB; AAF55246.
XX
PT New polynucleotides encoding mouse parkin2 protein, useful for producing
PT a transgenic non-human animal as an animal model for neurodegenerative
PT diseases.
XX
PS Claim 7; Page 21-22; 62pp; English.
XX
CC The present sequence represents a murine parkin2 polypeptide. Mutations
CC or deletions in the parkin2 gene cause Parkinson's disease in humans. The
CC human parkin2 gene is located in gene region 6q25.2-27. Parkin2
CC polypeptides and polynucleotides are useful for analysing
CC neurodegenerative diseases. They are also useful for testing the efficacy
CC of the treatment of a neurodegenerative disease such as Parkinson's
CC disease, Alzheimer's disease, Huntington's disease, amyotrophic lateral

CC sclerosis, Multi-system atrophy, Wilson's disease, Pick's disease, Prion
CC disease, and secondary causes inducing Parkinson's syndromes like toxins,
CC drugs, brain tumours, head trauma, stroke, vascular irregularities or
CC metabolic irregularities, associated with a less active or non-active
CC parkin protein
XX
SQ Sequence 250 AA;

Query Match 38.6%; Score 1002; DB 4; Length 250;
Best Local Similarity 77.6%; Pred. No. 1.7e-79;
Matches 190; Conservative 20; Mismatches 35; Indels 0; Gaps 0;

```
Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||:|||||||||||| ||||||||||||||||||||||| |  ||||||
Db          1 MIVFVRFNSSYGFPVEVDSDTSILQLKEVVAKRQGVPADQLRVIFAGKELPNHLTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              |:|||||||||||  |:  | ||:|||:|:: : |    | :|||||||||  || ||||||
Db         61 LEQQSIVHIVQRPRRRSHETNASGGDEPQSTSEGSIWESRSLTRVDLSSHTLPVDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||| |||::||  |  || :  ||||::|||||| :|||||||||| ||:|||||| |||
Db        121 VILDTDSKRDSEAARGPAVKPTYNSFFIYCKGPCHKVQPGKLRVQCGTCKQATLTLAQGP 180

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              ||||||||||||||||||| |||| ||||||||||||||:| |||:|| :| |:| ||
Db        181 SCWDDVLIPNRMSGECQSPDCPGTRAEFFFKCGAHPTSDKDTSVALNLITSNRRSIPCIA 240

Qy        241 CTDVR 245
              |||||
Db        241 CTDVR 245
```

Search completed: February 23, 2005, 13:50:12
Job time : 105.258 secs

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:36:20 ; Search time 25.8801 Seconds
(without alignments)
1341.256 Million cell updates/sec

Title: US-10-622-817-8
Perfect score: 2596
Sequence: 1 MIVFVRFNSSHGFPVEVDSD.....CWNCGCEWNRVCMGDHWFDV 465

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2596	100.0	465	4	US-09-601-844B-2	Sequence 2, Appli
2	2596	100.0	465	4	US-09-949-016-6516	Sequence 6516, Ap
3	2405	92.6	437	4	US-09-601-844B-4	Sequence 4, Appli
4	2405	92.6	437	4	US-09-949-016-6853	Sequence 6853, Ap
5	1696.5	65.4	316	4	US-09-949-016-6854	Sequence 6854, Ap
6	370.5	14.3	117	4	US-09-270-767-32183	Sequence 32183, A
7	240.5	9.3	503	4	US-09-914-259-13	Sequence 13, Appl
8	222	8.6	474	4	US-09-354-221-2	Sequence 2, Appli
9	222	8.6	487	4	US-09-949-016-10111	Sequence 10111, A
10	191	7.4	506	4	US-09-248-796A-15410	Sequence 15410, A
11	167	6.4	587	1	US-08-398-008A-2	Sequence 2, Appli
12	167	6.4	587	2	US-08-893-333-2	Sequence 2, Appli
13	149	5.7	464	4	US-09-538-092-598	Sequence 598, App
14	141	5.4	328	4	US-09-248-796A-16471	Sequence 16471, A
15	133	5.1	352	3	US-08-854-764-2	Sequence 2, Appli
16	133	5.1	352	5	PCT-US95-09377-2	Sequence 2, Appli
17	131.5	5.1	1106	4	US-09-949-016-9626	Sequence 9626, Ap
18	130.5	5.0	1242	4	US-09-488-270A-2	Sequence 2, Appli
19	128.5	4.9	229	2	US-08-726-306A-23	Sequence 23, Appl
20	128.5	4.9	229	3	US-08-840-146-20	Sequence 20, Appl
21	128.5	4.9	229	3	US-09-360-220-20	Sequence 20, Appl
22	128	4.9	160	4	US-09-370-838-205	Sequence 205, App
23	128	4.9	160	4	US-09-854-133-205	Sequence 205, App
24	127	4.9	2321	4	US-09-230-652-2	Sequence 2, Appli
25	126	4.9	237	4	US-09-248-796A-19062	Sequence 19062, A
26	126	4.9	846	1	US-08-149-103-3	Sequence 3, Appli
27	126	4.9	846	1	US-08-451-883-3	Sequence 3, Appli
28	126	4.9	873	1	US-08-393-734-2	Sequence 2, Appli
29	126	4.9	873	3	US-08-894-489-2	Sequence 2, Appli
30	126	4.9	904	4	US-09-949-016-9528	Sequence 9528, Ap
31	125.5	4.8	103	2	US-08-771-201-9	Sequence 9, Appli
32	125	4.8	76	1	US-08-232-815-2	Sequence 2, Appli
33	125	4.8	76	1	US-08-350-906-2	Sequence 2, Appli
34	125	4.8	76	4	US-09-601-844B-61	Sequence 61, Appl
35	125	4.8	76	5	PCT-US95-04536-2	Sequence 2, Appli
36	125	4.8	206	4	US-09-248-796A-19430	Sequence 19430, A
37	125	4.8	1121	1	US-07-789-915A-2	Sequence 2, Appli
38	125	4.8	1121	1	US-08-005-002C-2	Sequence 2, Appli
39	125	4.8	1121	1	US-08-487-203A-2	Sequence 2, Appli
40	124.5	4.8	77	6	5510474-3	Patent No. 5510474
41	124.5	4.8	77	6	5510474-3	Patent No. 5510474
42	124.5	4.8	305	4	US-09-866-153-6	Sequence 6, Appli
43	124.5	4.8	305	4	US-09-693-467A-6	Sequence 6, Appli
44	124.5	4.8	305	4	US-09-270-976-6	Sequence 6, Appli
45	124.5	4.8	381	4	US-09-866-153-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-09-601-844B-2
; Sequence 2, Application US/09601844B
; Patent No. 6716621
; GENERAL INFORMATION:
; APPLICANT: Shimizu, No. 6716621uyoshi
; APPLICANT: Mizuno, Yoshikuni
; TITLE OF INVENTION: Isolated DNA or Gene Responsible for Parkinson's Disease
; FILE REFERENCE: 0652.2110000
; CURRENT APPLICATION NUMBER: US/09/601,844B
; CURRENT FILING DATE: 2000-08-09
; PRIOR APPLICATION NUMBER: PCT/JP99/00545
; PRIOR FILING DATE: 1999-02-09
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-601-844B-2

Query Match 100.0%; Score 2596; DB 4; Length 465;
Best Local Similarity 100.0%; Pred. No. 8.4e-251;
Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
```

```
Db        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |||||||||||||||||||||||||||||||||||||||||||||
Db        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465


RESULT 2
US-09-949-016-6516
; Sequence 6516, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6516
;   LENGTH: 465
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-6516

  Query Match           100.0%;  Score 2596;  DB 4;  Length 465;
  Best Local Similarity  100.0%;  Pred. No. 8.4e-251;
  Matches  465;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |||||||||||||||||||||||||||||||||||||||||||||
Db        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465


RESULT 3
US-09-601-844B-4
, Sequence 4, Application US/09601844B
; Patent No. 6716621
; GENERAL INFORMATION:
;  APPLICANT: Shimizu, No. 6716621uyoshi
;  APPLICANT:  Mizuno, Yoshikuni
;  TITLE OF INVENTION: Isolated DNA or Gene Responsible for Parkinson's Disease
;  FILE REFERENCE: 0652.2110000
;  CURRENT APPLICATION NUMBER: US/09/601,844B
;  CURRENT FILING DATE:  2000-08-09
;  PRIOR APPLICATION NUMBER: PCT/JP99/00545
;  PRIOR FILING DATE: 1999-02-09
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 437
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-601-844B-4

  Query Match           92.6%;  Score 2405;  DB 4;  Length 437;
  Best Local Similarity  94.0%;  Pred. No. 9.6e-232;
  Matches  437;  Conservative    0;  Mismatches    0;  Indels   28;  Gaps    1;

Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQ-- 178

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
                                  ||||||||||||||||||||||||||||||||||||||
Db        179 --------------------------EFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 212

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 272

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 332

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 392

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |||||||||||||||||||||||||||||||||||||||||||||
Db        393 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 437


RESULT 4
US-09-949-016-6853
; Sequence 6853, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
```

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6853
;   LENGTH: 437
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-6853

  Query Match             92.6%;  Score 2405;  DB 4;  Length 437;
  Best Local Similarity   94.0%;  Pred. No. 9.6e-232;
  Matches  437;  Conservative    0;  Mismatches    0;  Indels   28;  Gaps    1;

Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQ-- 178

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
                                  ||||||||||||||||||||||||||||||||
Db        179 -------------------------EFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 212

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 272

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 332

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 392

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |||||||||||||||||||||||||||||||||||||||||||||
Db        393 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 437


RESULT 5
US-09-949-016-6854
; Sequence 6854, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6854
;   LENGTH: 316
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-6854

  Query Match             65.4%;  Score 1696.5;  DB 4;  Length 316;
  Best Local Similarity   68.0%;  Pred. No. 4.2e-161;
  Matches  316;  Conservative    0;  Mismatches    0;  Indels  149;  Gaps    1;

Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQ--- 57

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 ------------------------------------------------------------ 57

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 ------------------------------------------------------------ 57

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
                                  ||||||||||||||||||||||||||||||||
Db         58 -------------------------EFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 91

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         92 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 151

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 211

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 271

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |||||||||||||||||||||||||||||||||||||||||||||
Db        272 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 316


RESULT 6
US-09-270-767-32183
; Sequence 32183, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32183
;   LENGTH: 117
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-270-767-32183

  Query Match             14.3%;  Score 370.5;  DB 4;  Length 117;
  Best Local Similarity   53.3%;  Pred. No. 3.8e-29;
  Matches   64;  Conservative   11;  Mismatches   40;  Indels    5;  Gaps    2;

Qy        345 PDQRKVTCEGGNGLGCGFAFCRECKEAYHEGEC-SAVFEASGTTTQAYRVDERAAEQARW 403
              || |||||:    |||: ||| | : || |||      || | :  | ||   | :|||
Db          1 PDCRKVTCKN----GCGYVFCRNCLQGYHIGECLPEGTGASATNSCEYTVDPNRAAEARW 56

Qy        404 EAASKETIKKTTKPCPRCHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWF 463
              : ||  ||| :||||||:|  | |::|||||| | :  |  |||| |  || | ||| |||
Db         57 DEASNVTIKVSTKPCPKCRTPTERDGGCMHMVCTRAGCGFEWCWVCQTEWTRDCMGAHWF 116


RESULT 7
US-09-914-259-13
; Sequence 13, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
;  APPLICANT: Makowski, Lee
;  APPLICANT:  Hyman, Paul
;  APPLICANT:  Williams, Mark
;  TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
```

```
;  FILE REFERENCE: 8471-010-999
;  CURRENT APPLICATION NUMBER: US/09/914,259
;  CURRENT FILING DATE:  2000-11-21
;  NUMBER OF SEQ ID NOS: 180
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
;   LENGTH: 503
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-914-259-13

  Query Match             9.3%;  Score 240.5;  DB 4;  Length 503;
  Best Local Similarity   27.5%;  Pred. No. 3e-15;
  Matches   72;  Conservative   42;  Mismatches  111;  Indels   37;  Gaps   13;

Qy        209 FFKCGAHPTSDKETPVALHLIATNSRNITCITCTDVRSPVLVFQCNSRHVICLDCFHLYC 268
              |||| ||  :      |:    : |:   |  |      |  :      |  |: |:| |
Db        105 FFKC-AHVINPFNATEAIKQKTSRSQCEECEICFSQLPPDSMAGLECGHRFCMPCWHEYL 163

Qy        269 VTRLNDRQFVHDPQLGYSLPCVA-GCPNSLIKELHHFRILGEEQYN-RYQQYGAEECVLQ 326
              |:      :    || :: | | || : |: ::    :: : :   :|||       | :
Db        164 STK------IVAEGLGQTISCAAHGC-DILVDDVTVANLVTDARVRVKYQQLITNSFV-E 215

Qy        327 MGGVL--CPRPGCG-AGLLPEPDQRKVTCEGGNGLGCGFAFCRECKEAYHEG-ECSAVFE 382
              :|  ||   |  |  :|  : |:| |:       ||  ||  | | :|:  :|  :
Db        216 CNQLLRWCPSVDCTYAVKVPYAEPRRVHCK------CGHVFCFACGENWHDPVKCRWL-- 267

Qy        383 ASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPRCHVPVEKNGGCMHMKCPQPQCR 442
              :  :  :  :| : | ||:       || |||| | :||:||| || |    |:
Db        268 -----KKWIKKCDDDSETSNWIAAN-------TKECPRCSVTIEKDGGCNHMVCKNQNCK 315

Qy        443 LEWCWNCGCEWNRVCMGDHWFD 464
              |:|| |   |     |  |::
Db        316 NEFCWVCLGSWEP--HGSSWYN 335


RESULT 8
US-09-354-221-2
; Sequence 2, Application US/09354221
; Patent No. 6699714
; GENERAL INFORMATION:
;  APPLICANT: Chang, Chawnshang
;  TITLE OF INVENTION: Androgen Receptor Coactivators
;  FILE REFERENCE: 920920.90011
;  CURRENT APPLICATION NUMBER: US/09/354,221
;  CURRENT FILING DATE: 1999-07-15
;  EARLIER APPLICATION NUMBER: US 60/100,243
;  EARLIER FILING DATE: 1998-09-14
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 474
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-354-221-2

  Query Match             8.6%;  Score 222;  DB 4;  Length 474;
  Best Local Similarity   27.8%;  Pred. No. 2e-13;
  Matches   71;  Conservative   27;  Mismatches   97;  Indels   60;  Gaps   13;

Qy        232 NSRNITC--ITCTDVRSPVLVF-QCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLP 288
              ||:    |   |  : | :|  ||| |  |    |  :: | |             :
Db        214 NSKLFLCSICFCEKLGSECMYFLEC--RHVYCKACLKDYFEIQIRDGQ----------VQ 261

Qy        289 CVAGCPN---------SLIKELHHFRILGEEQYNRYQQYGAEECVLQMGGVL-CPRPGCG 338
              |:  ||            :|||     :  | : || :   :  :   |  |: |||| |
Db        262 CL-NCPEPKCPSVATPGQVKEL-----VEAELFARYDRLLLQSSLDLMADVVYCPRPCCQ 315

Qy        339 AGLLPEPDQRKVTCEGGNGLGCGFAFCRECKEAYH-EGECSAVFEASGTTTQAY------ 391
              :: ||       |  |    | ||||  |:  ||    |     |        |
Db        316 LPVMQEPG-----CTMGICSSCNFAFCTLCRLTYHGVSPCKVTAEKLMDLRNEYLQADEA 370

Qy        392 -------RVDERAAEQARWEAASKETIKKTTKPCPRCHVPVEKNGGCMHMKCPQPQCRLE 444
                    |  :|  ::|  |  ||| ::| :| || |  |:||  ||  | |    |
Db        371 NKRLLDQRYGKRVIQKALEEMESKEWLEKNSKSCPCCGTPIEKLDGCNKMTC--TGCMQY 428

Qy        445 WCWNCGCEWNRVCMG 459
              :||       :|||
Db        429 FCW--------ICMG 435


RESULT 9
US-09-949-016-10111
; Sequence 10111, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10111
;   LENGTH: 487
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-10111

  Query Match             8.6%;  Score 222;  DB 4;  Length 487;
  Best Local Similarity   27.8%;  Pred. No. 2e-13;
  Matches   71;  Conservative   27;  Mismatches   97;  Indels   60;  Gaps   13;

Qy        232 NSRNITC--ITCTDVRSPVLVF-QCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLP 288
              ||:    |   |  : | :|  ||| |  |    |  :: | |             :
Db        227 NSKLFLCSICFCEKLGSECMYFLEC--RHVYCKACLKDYFEIQIRDGQ----------VQ 274

Qy        289 CVAGCPN---------SLIKELHHFRILGEEQYNRYQQYGAEECVLQMGGVL-CPRPGCG 338
              |:  ||            :|||     :  | : || :   :  :   |  |: |||| |
Db        275 CL-NCPEPKCPSVATPGQVKEL-----VEAELFARYDRLLLQSSLDLMADVVYCPRPCCQ 328

Qy        339 AGLLPEPDQRKVTCEGGNGLGCGFAFCRECKEAYH-EGECSAVFEASGTTTQAY------ 391
              :: ||       |  |    | ||||  |:  ||    |     |        |
Db        329 LPVMQEPG-----CTMGICSSCNFAFCTLCRLTYHGVSPCKVTAEKLMDLRNEYLQADEA 383

Qy        392 -------RVDERAAEQARWEAASKETIKKTTKPCPRCHVPVEKNGGCMHMKCPQPQCRLE 444
                    |  :|  ::|  |  ||| ::| :| || |  |:||  ||  | |    |
Db        384 NKRLLDQRYGKRVIQKALEEMESKEWLEKNSKSCPCCGTPIEKLDGCNKMTC--TGCMQY 441

Qy        445 WCWNCGCEWNRVCMG 459
              :||       :|||
Db        442 FCW--------ICMG 448


RESULT 10
US-09-248-796A-15410
; Sequence 15410, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
```

```
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15410
;   LENGTH: 506
;   TYPE: PRT
;   ORGANISM: Candida albicans
US-09-248-796A-15410

  Query Match             7.4%;  Score 191;  DB 4;  Length 506;
  Best Local Similarity   23.6%;  Pred. No. 2.7e-10;
  Matches   79;  Conservative   48;  Mismatches  118;  Indels   90;  Gaps   17;

Qy        166 CSTCRQATLTLTQGPSCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVA 225
              |  | | ||   :| |                |     |    ||:|  |
Db        207 CDIC-QKTL---KGLHC--------TKFDSCGHVFCNNCLAEYFESC------------- 241

Qy        226 LHLIATNSRNITCITCTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGY 285
                    | :|  : | |        |:|  ::|     | :   | ::: : | |
Db        242 -----IESGDIDKVHCPD-------FECTKKYVDKKKKF-MELETWVHNDKKVKDMLNQL 288

Qy        286 SLPCVAGCPNSLIKELHHFRILGEEQYNRYQ--QYGAEECVLQMGGVLCPRPGCGAGLLP 343
                :| :   | ::: ::     | :  || ::  ||     :|    | ||| ||   :
Db        289 LIPSI---PLAMLTKILKSPSLVDRYYNLFKKSQYEWIGNLLPNRLVKCPRVGCDEVIFR 345

Qy        344 EP-DQRKVTCEGGNGLGCGFAFCRECKEAYHE--------GECSAVFEASGTTTQAY--- 391
              |  |:: | |       | :||| :|:::||        |              :||
Db        346 EDIDEKLVVCP-----KCKYAFCNDCRKSYHARFKLCLKINENDKYLGIPIEDLEAYPLL 400

Qy        392 ----------------RVDERAAEQARWEAASKETI--KKTTKPCPRCHVPVEKNGGCM 432
                             |:| || |: : :   :: :  :||   || | |  ||: ||
Db        401 PPDSYDKKIMNAKYGRKRID-RAIEEYQMDQLFQQMMRERKTVVQCPGCSVATEKSEGCN 459

Qy        433 HMKCPQPQCRLEWCWNCGCEWNRVCMG---DHWFD 464
              |||    |: ::|:||| :     :|   ||: |
Db        460 KMKC--SLCKTDFCFNCGSK-----IGNNHDHFVD 487


RESULT 11
US-08-398-008A-2
; Sequence 2, Application US/08398008A
; Patent No. 5665588
;  GENERAL INFORMATION:
;    APPLICANT:  Kornbluth, Jacki
;    TITLE OF INVENTION:  DNA Encoding Natural Killer Lytic Associated
;    TITLE OF INVENTION:  Protein
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Gilbreth & Adler, P.C.
;      STREET:  8011 Candle Lane
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 1.44 Mb storage
;      COMPUTER:  MACINTOSH IIci
;      OPERATING SYSTEM:  Macintosh
;      SOFTWARE:  Microsoft Word 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/398,008A
;      FILING DATE:  March 2, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/126,501
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Adler, Dr. Benjamin Aaron
;      REGISTRATION NUMBER:  35,423
;      REFERENCE/DOCKET NUMBER:  D5705CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (713) 777-2321
;      TELEFAX:  (713) 777-6908
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  587
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Protein
;    HYPOTHETICAL:  No
;    ANTI-SENSE:  No
US-08-398-008A-2

  Query Match             6.4%;  Score 167;  DB 1;  Length 587;
  Best Local Similarity   24.2%;  Pred. No. 8.4e-08;
  Matches   80;  Conservative   24;  Mismatches  114;  Indels  112;  Gaps   20;

Qy        178 QGPSCWDDVLIPNRMSGECQSPHCPGTSAEF-FFKCGAHPTSDKETPVALHLIATNSRNI 236
              |||          |  : |       |   ||      || | | |   |          :
Db         73 QGPP-------PEALPAE------PAAEAEAEAAAAGAEPGFDDEE--AAEGGGPGGEEV 117

Qy        237 TCITC----TDVRSPVLVFQCNSRHVICLDCFHLY---------------CVTRLND-- 274
              |  |       |:| |:   :  |  | ||   |                   |||
Db        118 ECPLCLVRLPPERAPRLL---SCPHRSCRDCLRHYLRLEISESRVPISCPECSERLNPHD 174

Qy        275 -RQFVHDPQLGYSLPCVAGCPNSLIKELHHFRILGEEQYNRYQQYGAEECVLQMGGVLCP 333
               |  : || |              :| :    ||   |       :|          ||
Db        175 IRLLLADPPL----------------MHKY----EEFIVRRYLASDPDC------RWCP 207

Qy        334 RPGCGAGLL-------PEPDQRKVTCEGGNGLGCGFAFCRECKEAYHEGECSAVFEASGT 386
               | ||  ::       |     |:|||     ||   ||  ||: :|  :   :  |
Db        208 APDCGYAVIAYGCASCP-----KLTCERE---GCQTEFCYHCKQIWHPNQTCDM--ARQQ 257

Qy        387 TTQAYRVDERAAEQARWEAASKETIKKTTKPCPRCHVPVEK--NGGCMHMKCPQPQCRLE 444
               |  ||  :      :    :|:     ||||||   : |  :| | || |    |  |
Db        258 RAQTLRVRTKHTSGLSY---GQESGPDDIKPCPRCSAYIIKMNDGSCNHMTC--AVCGCE 312

Qy        445 WCWNC-------------GC------EWNR 455
              :|| |             ||       |:|
Db        313 FCWLCMKEISDLHYLSPSGCTFWGKKPWSR 342


RESULT 12
US-08-893-333-2
; Sequence 2, Application US/08893333A
; Patent No. 5981705
; GENERAL INFORMATION:
;  APPLICANT: Kornbluth, Jacki
;  TITLE OF INVENTION: DNA Encoding Natural Killer Lytic Associated Protein
;  FILE REFERENCE: D5705CIP/D
;  CURRENT APPLICATION NUMBER: US/08/893,333A
;  CURRENT FILING DATE: 1997-07-16
;  NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 2
;   LENGTH: 587
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Amino acid sequence of Natural Killer Lytic
;   OTHER INFORMATION: Associated Protein encoded by nucleotides
;   OTHER INFORMATION: 190 to 1953 of Sequence ID. No. 5981705 1
; Patent No. 5981705
US-08-893-333-2

  Query Match             6.4%;  Score 167;  DB 2;  Length 587;
  Best Local Similarity   24.2%;  Pred. No. 8.4e-08;
  Matches   80;  Conservative   24;  Mismatches  114;  Indels  112;  Gaps   20;

Qy        178 QGPSCWDDVLIPNRMSGECQSPHCPGTSAEF-FFKCGAHPTSDKETPVALHLIATNSRNI 236
              |||          |  : |       |   ||      || | | |   |          :
Db         73 QGPP-------PEALPAE------PAAEAEAEAAAAGAEPGFDDEE--AAEGGGPGGEEV 117

Qy        237 TCITC----TDVRSPVLVFQCNSRHVICLDCFHLY---------------CVTRLND-- 274
              |  |       |:| |:   :  |  | ||   |                   |||
Db        118 ECPLCLVRLPPERAPRLL---SCPHRSCRDCLRHYLRLEISESRVPISCPECSERLNPHD 174
```

```
Qy        275 -RQFVHDPQLGYSLPCVAGCPNSLIKELHHFRILGEEQYNRYQQYGAEECVLQMGGVLCP 333
Db        175 IRLLLADPPL----------------MHKY----EEFIVRRYLASDPDC------RWCP 207

Qy        334 RPGCGAGLL-------PEPDQRKVTCEGGNGLGCGFAFCRECKEAYHEGECSAVFEASGT 386
Db        208 APDCGYAVIAYGCASCP-----KLTCERE---GCQTEFCYHCKQIWHPNQTCDM--ARQQ 257

Qy        387 TTQAYRVDERAAEQARWEAASKETIKKTTKPCPRCHVPVEK--NGGCMHMKCPQPQCRLE 444
Db        258 RAQTLRVRTKHTSGLSY---GQESGPDDIKPCPRCSAYIIKMNDGSCNHMTC--AVCGCE 312

Qy        445 WCWNC-------------GC------EWNR 455
Db        313 FCWLCMKEISDLHYLSPSGCTFWGKKPWSR 342
```

RESULT 13
US-09-538-092-598
; Sequence 598, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 598
; LENGTH: 464
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number YML068W
US-09-538-092-598

Query Match 5.7%; Score 149; DB 4; Length 464;
Best Local Similarity 21.5%; Pred. No. 3.8e-06;
Matches 59; Conservative 33; Mismatches 105; Indels 78; Gaps 13;

```
Qy        235 NITCITCTDVRSPVLVFQ-----CNSRHVICLDCFHLYCVTRLNDRQF--VHDPQLGYS- 286
Db        177 NYHCCICMEMEKGVRMIKLPCENANVEHYLCRGCAKSYFTAMIQENRISSVRCPQCEYKE 236

Qy        287 ----------------LPCVAGCPNSLIKELHHFRILGEEQYNRYQQYGAEECVLQMGG 329
Db        237 LKLEDFKSYKKMLKALFTPLI---PVSFLKE-----VIDTELCERYEKMFYNQAATRLSK 288

Qy        330 ------VLCPRPGCGAGLLPEP-DQRKVTCEGGNGLGCGFAFCRECKEAYH--EGECSAV 380
Db        289 YCPYACVTCRR--CDSWCTKEDLDDAMIQCQ-----KCHFVFCFDCLHAWHGYNNKCGKK 341

Qy        381 FEASGTTTQAY------------------------RVDERAAEQARWEAASKETIKKTT 415
Db        342 VSLSTDIIEEYLDDTVTSYERKRKLEAKYGRRVLELEVNDYLAEKMLDLAIKKE--GSNL 399

Qy        416 KPCPRCHVPVEKNGGCMHMKCPQPQCRLEWCWNCG 450
Db        400 QRCPKCKVVVERSEGCNKMKC--EVCGTLFCFICG 432
```

RESULT 14
US-09-248-796A-16471
; Sequence 16471, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16471
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-16471

Query Match 5.4%; Score 141; DB 4; Length 328;
Best Local Similarity 30.2%; Pred. No. 1.4e-05;
Matches 32; Conservative 15; Mismatches 37; Indels 22; Gaps 5;

```
Qy        358 LGCG--FAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTT 415
Db         58 IGCSEHHEFCFNCKYENHL-PCPC-----------WVSKRWIKKCNDDSETAHWIDANT 104

Qy        416 KPCPRCHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDH 461
Db        105 HSCPKCYSSIEKNGGCNHMTC--QKCKYEFCWVCLKDWS-----DH 143
```

RESULT 15
US-08-854-764-2
; Sequence 2, Application US/08854764
; Patent No. 6103500
; GENERAL INFORMATION:
; APPLICANT: Innes, Michael
; APPLICANT: Creasey, Abla
; TITLE OF INVENTION: Production of Tissue Factor Pathway
; TITLE OF INVENTION: Inhibitor
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton St.
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,764
; FILING DATE: 12-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/286,530
; FILING DATE: 05-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Savereide, Paul B.
; REGISTRATION NUMBER: 36,914
; REFERENCE/DOCKET NUMBER: 0991.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2585
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-854-764-2

Query Match 5.1%; Score 133; DB 3; Length 352;
Best Local Similarity 18.3%; Pred. No. 0.0001;
Matches 70; Conservative 56; Mismatches 116; Indels 140; Gaps 16;

```
Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              | :||:  :     :||:|  :|  :|  :  ::|:| || |:|||||:| :  |: : :
Db          1 MQIFVKTLTGKTITLEVESSDTIDNVKSKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYN 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              : ::| :|:| |            |||         : |   :|  :|    |
Db         61 IQKESTLHLVLR---------LRGGDSEE------DEEHTIITDTELPPLKL-------- 97

Qy        121 VILHT--DSRKDSPPAGSPAGRSIYNSF------YVY---------------CKGPCQRV 157
               :|:    : |  |  :   |  :| |       ::|               ||  | |
Db         98 --MHSFCAFKADDGPCKAIMKRFFFNIFTRQCEEFIYGGCEGNQNRFESLEECKKMCTRD 155

Qy        158 QPGKLRVQCSTCRQATLTLTQGPSCWDDVLIPNRMSGECQSPH-CPGTSAEFFFK----- 211
              ::        : ||   :   |: :           : |  | |    :|:
Db        156 NANRI-------IKTTLQQEKPDFCFLE----------EDPGICRGYITRYFYNNQTKQ 197

Qy        212 -------------------------CGAHPTSDKETPVALHLIATNSRNITCITCTDVRS 246
                                       |   |   :       | | |:     :|    :
Db        198 CERFKYGGCLGNMNNFETLEECKNICEDGPNGFQVDNYGTQLNAVNN----SLTPQSTKV 253

Qy        247 PVLVFQCNSRHVICLDCFH--LYCVT-------RLNDRQFVHDPQLGYSLP--------- 288
              | | |:           ||   :|:|       | |: :| ::  :|   |
Db        254 PSL-FE----------FHGPSWCLTPADRGLCRANENRFYYNSVIGKCRPFKYSGCGGN 301

Qy        289 ---------CVAGCPNSLIKEL 301
                     |:  |    |: :
Db        302 ENNFTSKQECLRACKKGFIQRI 323
```

Search completed: February 23, 2005, 14:01:03
Job time : 27.8801 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:57:26 ; Search time 73.013 Seconds
(without alignments)
2084.158 Million cell updates/sec

Title: US-10-622-817-8
Perfect score: 2596
Sequence: 1 MIVFVRFNSSHGFPVEVDSD.....CWNCGCEWNRVCMGDHWFDV 465

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1380268 seqs, 327241040 residues

Total number of hits satisfying chosen parameters: 1380268

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2596	100.0	465	16	US-10-473-226-2	Sequence 2, Appli
2	2596	100.0	465	16	US-10-776-604-2	Sequence 2, Appli
3	2596	100.0	465	17	US-10-839-688-9	Sequence 9, Appli
4	2584	99.5	465	16	US-10-473-226-4	Sequence 4, Appli
5	2405	92.6	437	16	US-10-776-604-4	Sequence 4, Appli
6	870	33.5	156	9	US-09-785-548-4	Sequence 4, Appli
7	544.5	21.0	386	14	US-10-239-249-2	Sequence 2, Appli
8	373	14.4	64	15	US-10-313-203-10	Sequence 10, Appl
9	323	12.4	56	15	US-10-313-203-17	Sequence 17, Appl
10	263	10.1	46	9	US-09-864-761-36750	Sequence 36750, A
11	246	9.5	525	16	US-10-437-963-199787	Sequence 199787,
12	242.5	9.3	604	15	US-10-425-114-55407	Sequence 55407, A
13	240.5	9.3	503	14	US-10-080-608A-13	Sequence 13, Appl
14	240.5	9.3	503	15	US-10-370-685-102	Sequence 102, App
15	233.5	9.0	585	15	US-10-424-599-147673	Sequence 147673,
16	224.5	8.6	648	15	US-10-425-114-63379	Sequence 63379, A
17	222	8.6	474	9	US-09-808-387-42	Sequence 42, Appl
18	222	8.6	474	14	US-10-442-754-2	Sequence 2, Appli
19	222	8.6	474	16	US-10-408-765A-572	Sequence 572, App
20	222	8.6	477	15	US-10-264-049-3019	Sequence 3019, Ap
21	219.5	8.5	1753	14	US-10-146-473-44	Sequence 44, Appl
22	219.5	8.5	1753	15	US-10-276-774-1824	Sequence 1824, Ap
23	219.5	8.5	2517	15	US-10-313-203-2	Sequence 2, Appli
24	215	8.3	32	15	US-10-313-203-18	Sequence 18, Appl
25	207.5	8.0	623	16	US-10-437-963-148905	Sequence 148905,
26	206.5	8.0	452	15	US-10-425-114-38443	Sequence 38443, A
27	205.5	7.9	449	15	US-10-425-114-39657	Sequence 39657, A
28	204	7.9	624	15	US-10-374-780A-1727	Sequence 1727, Ap
29	204	7.9	624	16	US-10-437-963-140793	Sequence 140793,
30	202.5	7.8	289	16	US-10-437-963-133385	Sequence 133385,
31	201	7.7	333	15	US-10-424-599-217796	Sequence 217796,
32	200	7.7	593	15	US-10-225-066A-854	Sequence 854, App
33	200	7.7	593	15	US-10-225-067-56	Sequence 56, Appl
34	200	7.7	593	15	US-10-374-780A-330	Sequence 330, App
35	195.5	7.5	541	15	US-10-424-599-265215	Sequence 265215,
36	193	7.4	606	16	US-10-437-963-111378	Sequence 111378,
37	192	7.4	354	15	US-10-425-114-56248	Sequence 56248, A
38	191.5	7.4	540	16	US-10-437-963-168788	Sequence 168788,
39	191	7.4	338	15	US-10-425-114-56467	Sequence 56467, A
40	184	7.1	583	16	US-10-437-963-124556	Sequence 124556,
41	182.5	7.0	522	16	US-10-437-963-184651	Sequence 184651,
42	180	6.9	250	16	US-10-437-963-152161	Sequence 152161,
43	180	6.9	394	15	US-10-425-114-54725	Sequence 54725, A
44	180	6.9	557	15	US-10-424-599-212592	Sequence 212592,
45	177.5	6.8	396	15	US-10-425-114-40231	Sequence 40231, A

ALIGNMENTS

RESULT 1
US-10-473-226-2
; Sequence 2, Application US/10473226
; Publication No. US20040198650A1
; GENERAL INFORMATION:
; APPLICANT: NsGene A/S
; TITLE OF INVENTION: Means for inhibiting proteolytical processing of Parkin
; FILE REFERENCE: 506-204-WO
; CURRENT APPLICATION NUMBER: US/10/473,226
; CURRENT FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: DK PA 2001 00525
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: US 60/281,286
; PRIOR FILING DATE: 2001-04-03
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (1)..()
; OTHER INFORMATION: Native Parkin
US-10-473-226-2

Query Match 100.0%; Score 2596; DB 16; Length 465;
Best Local Similarity 100.0%; Pred. No. 2.2e-220;
Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |||||||||||||||||||||||||||||||||||||||||||||
Db        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
```

RESULT 2
US-10-776-604-2
; Sequence 2, Application US/10776604
; Publication No. US20050003385A1
; GENERAL INFORMATION:
; APPLICANT: Shimizu, Nobuyoshi
; APPLICANT: Mizuno, Yoshikuni
; TITLE OF INVENTION: DNAs or Genes Participating in Parkinson's Disease
; FILE REFERENCE: 0652.2110001
; CURRENT APPLICATION NUMBER: US/10/776,604
; CURRENT FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 09/601,844
; PRIOR FILING DATE: 2000-08-09
; PRIOR APPLICATION NUMBER: PCT/JP99/00545
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: JP 10/27531
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-776-604-2

Query Match 100.0%; Score 2596; DB 16; Length 465;
Best Local Similarity 100.0%; Pred. No. 2.2e-220;
Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |||||||||||||||||||||||||||||||||||||||||||||
Db        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
```

RESULT 3
US-10-839-688-9
; Sequence 9, Application US/10839688
; Publication No. US20050014173A1
; GENERAL INFORMATION:
; APPLICANT: Farrer, Matthew J.
; TITLE OF INVENTION: PARKINSON'S DISEASE MARKERS
; FILE REFERENCE: 07039-448001
; CURRENT APPLICATION NUMBER: US/10/839,688
; CURRENT FILING DATE: 2004-05-05
; PRIOR APPLICATION NUMBER: US 60/468,832
; PRIOR FILING DATE: 2003-05-08
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-839-688-9

Query Match 100.0%; Score 2596; DB 17; Length 465;
Best Local Similarity 100.0%; Pred. No. 2.2e-220;
Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |||||||||||||||||||||||||||||||||||||||||||||
Db        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
```

```
RESULT 4
US-10-473-226-4
; Sequence 4, Application US/10473226
; Publication No. US20040198650A1
; GENERAL INFORMATION:
;  APPLICANT: NsGene A/S
;  TITLE OF INVENTION: Means for inhibiting proteolytical processing of Parkin
;  FILE REFERENCE: 506-204-WO
;  CURRENT APPLICATION NUMBER: US/10/473,226
;  CURRENT FILING DATE:  2003-09-29
;  PRIOR APPLICATION NUMBER: DK PA 2001 00525
;  PRIOR FILING DATE: 2001-03-29
;  PRIOR APPLICATION NUMBER: US 60/281,286
;  PRIOR FILING DATE: 2001-04-03
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 465
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: (1)..()
;   OTHER INFORMATION: Parkin with a D126E mutation
US-10-473-226-4

  Query Match             99.5%;  Score 2584;  DB 16;  Length 465;
  Best Local Similarity   99.6%;  Pred. No. 2.5e-219;
  Matches  463;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VILHTESRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETSVALHLIATNSRNITCIT 240

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |||||||||||||||||||||||||||||||||||||||||||||
Db        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465


RESULT 5
US-10-776-604-4
; Sequence 4, Application US/10776604
; Publication No. US20050003385A1
; GENERAL INFORMATION:
;  APPLICANT: Shimizu, Nobuyoshi
;  APPLICANT:  Mizuno, Yoshikuni
;  TITLE OF INVENTION: DNAs or Genes Participating in Parkinson's Disease
;  FILE REFERENCE: 0652.2110001
;  CURRENT APPLICATION NUMBER: US/10/776,604
;  CURRENT FILING DATE:  2004-02-12
;  PRIOR APPLICATION NUMBER: 09/601,844
;  PRIOR FILING DATE: 2000-08-09
;  PRIOR APPLICATION NUMBER: PCT/JP99/00545
;  PRIOR FILING DATE: 1999-02-09
;  PRIOR APPLICATION NUMBER: JP 10/27531
;  PRIOR FILING DATE: 1998-02-09
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 437
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-776-604-4

  Query Match             92.6%;  Score 2405;  DB 16;  Length 437;
  Best Local Similarity   94.0%;  Pred. No. 1.5e-203;
  Matches  437;  Conservative    0;  Mismatches    0;  Indels   28;  Gaps    1;

Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQ-- 178

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
                                          ||||||||||||||||||||||||||||||||
Db        179 -------------------------EFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 212

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 272

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 332

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 392

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |||||||||||||||||||||||||||||||||||||||||||||
Db        393 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 437


RESULT 6
US-09-785-548-4
; Sequence 4, Application US/09785548
; Patent No. US20020155577A1
; GENERAL INFORMATION:
;  APPLICANT: AVENTIS PHARMACEUTICALS, INC.
;  TITLE OF INVENTION: COMPOSITIONS THAT CAN BE USED FOR REGULATING THE ACTIVITY OF PARKI
;  FILE REFERENCE: ST00005
;  CURRENT APPLICATION NUMBER: US/09/785,548
;  CURRENT FILING DATE:  2001-02-20
;  NUMBER OF SEQ ID NOS: 46
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 156
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-785-548-4

  Query Match             33.5%;  Score 870;  DB 9;  Length 156;
  Best Local Similarity   99.4%;  Pred. No. 1.3e-68;
  Matches  155;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy        135 GSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGPSCWDDVLIPNRMSG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGPSCWDDVLIPNRMSG 60

Qy        195 ECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCITCTDVRSPVLVFQCN 254
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         61 ECQSPHCPGTSAEFFFKCGAHPTSDKETSVALHLIATNSRNITCITCTDVRSPVLVFQCN 120

Qy        255 SRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCV 290
              ||||||||||||||||||||||||||||||||||||
Db        121 SRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCV 156


RESULT 7
US-10-239-249-2
; Sequence 2, Application US/10239249.
; Publication No. US20030177507A1
; GENERAL INFORMATION:
;  APPLICANT: HONER, MARIUS
;  APPLICANT:  LINK, WOLFGANG
;  APPLICANT:  BAUMEISTER, RALF
;  TITLE OF INVENTION: NEMATODES AS MODEL ORGANISMS FOR INVESTIGATING
;  TITLE OF INVENTION:  NEURODEGERATIVE DISEASES AND, IN PARTICULAR, PARKINSON'S
;  TITLE OF INVENTION:  DISEASE, USES AND METHODS FOR FINDING SUBSTANCES AND
;  TITLE OF INVENTION:  GENES WHICH CAN BE USED IN TREATING SUCH DISEASES, AND
;  TITLE OF INVENTION:  IDENTIFICATION OF A NEMATODE GENE
;  FILE REFERENCE: 02481.1804-00000
;  CURRENT APPLICATION NUMBER: US/10/239,249
;  CURRENT FILING DATE:  2002-09-20
;  PRIOR APPLICATION NUMBER: PCT/EP01/03214
;  PRIOR FILING DATE: 2001-03-21
;  PRIOR APPLICATION NUMBER: 100 14 109.9
;  PRIOR FILING DATE: 2000-03-22
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 386
;   TYPE: PRT
;   ORGANISM: Caenorhabditis elegans
US-10-239-249-2

  Query Match             21.0%;  Score 544.5;  DB 14;  Length 386;
  Best Local Similarity   28.9%;  Pred. No. 2.1e-39;
  Matches  129;  Conservative   59;  Mismatches  166;  Indels   93;  Gaps   13;

Qy         22 SIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCDLDQQSIVHIVQRPWRKGQEMN 81
              :|  | : | |   :|:|:| |:| ||:|     :::  |   : : :::  :    |
Db         29 NIEDLTKDVEKLTEIPSDELEVVFCGKKLSKSTIMRDLSLTPATQIMLLRPKFNSHNENG 88

Qy         82 ATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLAVILHTDSRKDSPPAGSPAGRS 141
              ||                                  : |||              |
Db         89 AT-----------------------------------TAKITTDS------------S 99

Qy        142 IYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGPSCWDDVLIPNRMSGECQSPHC 201
              |  ||||:||  |  |: ||||| |  |   :: :   |  | |||   |:   |:
Db        100 ILGSFYVWCKN-CDDVKRGKLRVYCQKCSSTSVLVKSEPQNWSDVLKSKRIPAVCEECCT 158

Qy        202 PGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCITCTDVRSPVLVFQCNSRHVICL 261
              ||  ||| |||     :  :   ||  :  | :   |  |     :    ||  |: |
Db        159 PGLFAEFKFKC----LACNDPAAALTHVRGNWQMTECCVCDGKEKVIFDLGCN--HITCQ 212

Qy        262 DCFHLYCVTRLNDRQFVHDPQLGYSLPC-VAGCPNSLIKELHHFRILGEEQYNRYQQYGA 320
              ||  | :::|     ||: |  |::: |   || | :::::||| |:|:  |: ||:
Db        213 FCFRDYLLSQLERFGFVNQPPHGFTIFCPYPGC-NRVVQDVHHFHIMGQTSYSEYQRKAT 271

Qy        321 EECV-LQMGGVLCPRPGCGAGLLPEP--DQRKVTCEGGNGLGCGFAFCRECKEAYHEGEC 377
              |  : :   || ||   ||     ||  |  :  |       | |:|||:|     |  |
Db        272 ERLIAVDDKGVTCPNVSCGQSFFWEPYDDDGRSQCP-----DCFFSFCRKC----FERNC 322

Qy        378 SAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPRCHVPVEKNGGCMHMKCP 437
                 |   | |                          ||  ||: ||:|||  |:|||| |: |
Db        323 VCQSEDDLTRT--------------------TIDATTRRCPKCHVATERNGGCAHIHC- 360

Qy        438 QPQCRLEWCWNCGCEWNRVCMGDHWFD 464
                   | ::||: |  ||   |  ||||:
Db        361 -TSCGMDWCFKCKTEWKEECQWDHWFN 386


RESULT 8
US-10-313-203-10
; Sequence 10, Application US/10313203
; Publication No. US20040029134A1
; GENERAL INFORMATION:
;  APPLICANT: Gu, Wei
;  APPLICANT:  Nikolaev, Anatoly
;  TITLE OF INVENTION: P53-ASSOCIATED PARKIN-LIKE CYTOPLASMIC PROTEIN, AND RELATED COMPOS
;  TITLE OF INVENTION:  AND METHODS
;  FILE REFERENCE: 68106
;  CURRENT APPLICATION NUMBER: US/10/313,203
;  CURRENT FILING DATE:  2002-12-06
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
;   LENGTH: 64
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-313-203-10

  Query Match             14.4%;  Score 373;  DB 15;  Length 64;
  Best Local Similarity   100.0%;  Pred. No. 3e-25;
  Matches   64;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        314 RYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGCGFAFCRECKEAYH 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGCGFAFCRECKEAYH 60

Qy        374 EGEC 377
              ||||
Db         61 EGEC 64


RESULT 9
US-10-313-203-17
; Sequence 17, Application US/10313203
; Publication No. US20040029134A1
; GENERAL INFORMATION:
;  APPLICANT: Gu, Wei
;  APPLICANT:  Nikolaev, Anatoly
;  TITLE OF INVENTION: P53-ASSOCIATED PARKIN-LIKE CYTOPLASMIC PROTEIN, AND RELATED COMPOS
;  TITLE OF INVENTION:  AND METHODS
;  FILE REFERENCE: 68106
;  CURRENT APPLICATION NUMBER: US/10/313,203
;  CURRENT FILING DATE:  2002-12-06
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
;   LENGTH: 56
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-313-203-17

  Query Match             12.4%;  Score 323;  DB 15;  Length 56;
  Best Local Similarity   100.0%;  Pred. No. 6.7e-21;
  Matches   56;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        238 CITCTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGC 293
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CITCTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGC 56


RESULT 10
US-09-864-761-36750
; Sequence 36750, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 36750
;   LENGTH: 46
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL035697.18
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.68
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.76
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.94
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.63
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.68
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.91
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.77
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.73
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.96
;   OTHER INFORMATION: EST_HUMAN HIT: T78558.1, EVALUE 2.00e+00
;   OTHER INFORMATION: SWISSPROT HIT: Q04833, EVALUE 9.10e-01
US-09-864-761-36750

  Query Match             10.1%;  Score 263;  DB 9;  Length 46;
  Best Local Similarity   100.0%;  Pred. No. 1e-15;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        312 YNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNG 357
              |||||||||||||||||||||||||||||||||||||||||||||
Db          1 YNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNG 46
```

```
RESULT 11
US-10-437-963-199787
; Sequence 199787, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 199787
;   LENGTH: 525
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_95320C.1.pep
US-10-437-963-199787

  Query Match             9.5%;  Score 246;  DB 16;  Length 525;
  Best Local Similarity   30.1%;  Pred. No. 7.1e-13;
  Matches   71;  Conservative   26;  Mismatches   93;  Indels   46;  Gaps   11;

Qy        232 NSRNITCITCTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCV- 290
              | | :||  | ||     :      |  |: |:  |       |  | |     || |
Db        127 NRRRLTCAICFDVFDTGGMRSAGCSHFYCVSCWRGYV------RAAVGDGARCLSLRCPD 180

Qy        291 AGCPNSLIKELHHFRILGEEQYNRYQQYGAEECVLQMGGV-LCPRPGCGAGLLPEPDQRK 349
                || ::::||      ||::  |:  :     | :  |:  || |||          |
Db        181 PSCPAAVVRELVDAVADGEDR-ERFGWFALRSYVEESAGMRWCPGPGC---------SRA 230

Qy        350 VTCEGGNG--------LGCGFAFCREC-KEAYHEGECSAVFEASGTTTQAYRVDERAAEQ 400
              |  ||              ||   |  | :||:    |  |          |  |:: ::|
Db        231 VEFVGGGDGEESSEVFCSCGHGLCWRCGEEAHRPVSCKTV---------AKWVEKNSSES 281

Qy        401 --ARWEAASKETIKKTTKPCPRCHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWN 454
                | |  |         || ||:| :|:||| ||||| | :| |  |:|| |   |:
Db        282 ETATWLLAH-------TKHCPKCRLPIEKNLGCMHMTC-RPPCLHEFCWICLKPWH 329


RESULT 12
US-10-425-114-55407
; Sequence 55407, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 55407
;   LENGTH: 604
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3051-106-A7_FLI.pep
```

```
US-10-425-114-55407

  Query Match             9.3%;  Score 242.5;  DB 15;  Length 604;
  Best Local Similarity   27.7%;  Pred. No. 1.7e-12;
  Matches   70;  Conservative   31;  Mismatches   97;  Indels   55;  Gaps   12;

Qy        221 ETPVALHLIATNSRNITCITCTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQ---- 276
              | ||  :   :|:| :||  | ::      |     |  |  |:  |  | :||
Db        140 EKPVFQN---SNTRELTCGICFEMYPRARVESTACGHPYCYSCWAGYFSTSINDGPGCLV 196

Qy        277 -FVHDPQLGYSLPCVAGCPNSLIKELHHFRILGEEQYNRYQQYGAEECVLQMGGV-LCPR 334
               ||       | |     :|  |       :|   :| :|     :         ||
Db        197 LRCPDPS------CGAAVGQDMINLL-----ASDEDKQKYDRYLLRSYIEDNKKTKWCPA 245

Qy        335 PGCGAGLLPEPDQRKVTCEGGNG-----LGCGFAFCREC-KEAYHEGECSAVFEASGTTT 388
              |||         :  || : |:|       | ::||  | :||:   :|        || :
Db        246 PGC---------EYAVTFDAGSGNYDVSCLCSYSFCWNCTEEAHRPVDC-------GTVS 289

Qy        389 QAYRVDERAAEQARWEAASKETIKKTTKPCPRCHVPVEKNGGCMHMKCPQPQCRLEWCWN 448
              :   :  :|   |       |   :||||:|  |:||| ||||| |  | |: |:||
Db        290 KWILKNSAESENMNW-------ILANSKPCPKCKRPIEKNQGCMHMTC-TPPCKFEFCWL 341

Qy        449 CGCEWNRVCMGDH 461
              |   |:     ||
Db        342 CLGAWS-----DH 349


RESULT 13
US-10-080-608A-13
; Sequence 13, Application US/10080608A
; Publication No. US20030198956A1
; GENERAL INFORMATION:
;  APPLICANT: Makowski, Lee
;  APPLICANT:  Hyman, Paul
;  APPLICANT:  Williams, Mark
;  TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
;  FILE REFERENCE: 8471-010-999
;  CURRENT APPLICATION NUMBER: US/10/080,608A
;  CURRENT FILING DATE:  2002-02-21
;  NUMBER OF SEQ ID NOS: 180
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
;   LENGTH: 503
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-10-080-608A-13

  Query Match             9.3%;  Score 240.5;  DB 14;  Length 503;
  Best Local Similarity   27.5%;  Pred. No. 2.1e-12;
  Matches   72;  Conservative   42;  Mismatches  111;  Indels   37;  Gaps   13;

Qy        209 FFKCGAHPTSDKETPVALHLIATNSRNITCITCTDVRSPVLVFQCNSRHVICLDCFHLYC 268
              |||| ||  :       |:    : |:   |  |      |  :       |  |: |:| |
Db        105 FFKC-AHVINPFNATEAIKQKTSRSQCEECEICFSQLPPDSMAGLECGHRFCMPCWHEYL 163

Qy        269 VTRLNDRQFVHDPQLGYSLPCVA-GCPNSLIKELHHFRILGEEQYN-RYQQYGAEECVLQ 326
              |:       :    || :: | | || : |: ::    :: : :   :|||       | :
Db        164 STK------IVAEGLGQTISCAAHGC-DILVDDVTVANLVTDARVRVKYQQLITNSFV-E 215

Qy        327 MGGVL--CPRPGCG-AGLLPEPDQRKVTCEGGNGLGCGFAFCRECKEAYHEG-ECSAVFE 382
              :|  ||    |  |  :|  : |:| |:      ||  ||  | | :|:  :|  :
Db        216 CNQLLRWCPSVDCTYAVKVPYAEPRRVHCK------CGHVFCFACGENWHDPVKCRWL-- 267

Qy        383 ASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPRCHVPVEKNGGCMHMKCPQPQCR 442
                  :  :  :|  : | ||:       || |||| | :||  | |  |     |:
Db        268 -----KKWIKKCDDDSETSNWIAAN-------TKECPRCSVTIEKDGGCNHMVCKNQNCK 315

Qy        443 LEWCWNCGCEWNRVCMGDHWFD 464
              |:|| |   |      |  |::
Db        316 NEFCWVCLGSWEP--HGSSWYN 335


RESULT 14
```

```
US-10-370-685-102
; Sequence 102, Application US/10370685
; Publication No. US20030215903A1
; GENERAL INFORMATION:
;  APPLICANT: Hyman, Paul
;  APPLICANT:  Goldberg, Edward
;  TITLE OF INVENTION: Nanostructures Containing PNA Joining and Functional Elements
;  FILE REFERENCE: NANF.P-004
;  CURRENT APPLICATION NUMBER: US/10/370,685
;  CURRENT FILING DATE:  2003-02-21
;  PRIOR APPLICATION NUMBER: 10/080,608
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 159
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 102
;   LENGTH: 503
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-10-370-685-102

  Query Match             9.3%;  Score 240.5;  DB 15;  Length 503;
  Best Local Similarity   27.5%;  Pred. No. 2.1e-12;
  Matches   72;  Conservative   42;  Mismatches  111;  Indels   37;  Gaps   13;

Qy        209 FFKCGAHPTSDKETPVALHLIATNSRNITCITCTDVRSPVLVFQCNSRHVICLDCFHLYC 268
              |||| ||  :       |:    : |:   |  |      |  :       |  |: |:| |
Db        105 FFKC-AHVINPFNATEAIKQKTSRSQCEECEICFSQLPPDSMAGLECGHRFCMPCWHEYL 163

Qy        269 VTRLNDRQFVHDPQLGYSLPCVA-GCPNSLIKELHHFRILGEEQYN-RYQQYGAEECVLQ 326
              |:       :    || :: | | || : |: ::    :: : :   :|||       | :
Db        164 STK------IVAEGLGQTISCAAHGC-DILVDDVTVANLVTDARVRVKYQQLITNSFV-E 215

Qy        327 MGGVL--CPRPGCG-AGLLPEPDQRKVTCEGGNGLGCGFAFCRECKEAYHEG-ECSAVFE 382
              :|  ||    |  |  :|  : |:| |:      ||  ||  | | :|:  :|  :
Db        216 CNQLLRWCPSVDCTYAVKVPYAEPRRVHCK------CGHVFCFACGENWHDPVKCRWL-- 267

Qy        383 ASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPRCHVPVEKNGGCMHMKCPQPQCR 442
                  :  :  :|  : | ||:       || |||| | :||:||| || |    |:
Db        268 -----KKWIKKCDDDSETSNWIAAN-------TKECPRCSVTIEKDGGCNHMVCKNQNCK 315

Qy        443 LEWCWNCGCEWNRVCMGDHWFD 464
              |:|| |   |      |  |::
Db        316 NEFCWVCLGSWEP--HGSSWYN 335


RESULT 15
US-10-424-599-147673
; Sequence 147673, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 147673
;   LENGTH: 585
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(585)
;   OTHER INFORMATION: unsure at all Xaa locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_10436C.1.pep
US-10-424-599-147673

  Query Match             9.0%;  Score 233.5;  DB 15;  Length 585;
```

Best Local Similarity 26.5%; Pred. No. 1e-11;
Matches 67; Conservative 29; Mismatches 102; Indels 55; Gaps 12;

```
Qy        221 ETPVALHLIATNSRNITCITCTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLND----RQ 276
              | |:  |    |:| :||  | :      :   :  |   : |:  |  | :||
Db        120 EKPIVQH---PNARELTCGICFENYPRARIEMASCGHPYRISCWEGYISTSINDGPGCLM 176

Qy        277 FVHDPQLGYSLPCVAGCPNSLIKELHHFRILGEEQYNRYQQYGAEECVL-QMGGVLCPRP 335
                ||       | |    |       |    :|   :| :|     :         || |
Db        177 LSPDPX------CGAATGRS-----SRFGSRSDEDKQKYARYLLRSYIEDNKKSKWCPAP 225

Qy        336 GCGAGLLPEPDQRKVTCEGGN------GLGCGFAFCREC-KEAYHEGECSAVFEASGTTT 388
              ||         :  || : |:         | : ||  | :||:   :|        ||
Db        226 GC---------EYAVTFDAGSTGNYDVSCLCSYGFCWNCTEEAHRPVDC-------GTVA 269

Qy        389 QAYRVDERAAEQARWEAASKETIKKTTKPCPRCHVPVEKNGGCMHMKCPQPQCRLEWCWN 448
              :    :   :|   |       |   :||||:|  |:||| ||||| |  | |: |:||
Db        270 KWILKNSAESENMNW-------ILANSKPCPKCKRPIEKNHGCMHMTC-TPPCKFEFCWL 321

Qy        449 CGCEWNRVCMGDH 461
              |   |:     ||
Db        322 CVGAWS-----DH 329
```

Search completed: February 23, 2005, 14:33:24
Job time : 75.0113 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:34:50 ; Search time 19.3575 Seconds
(without alignments)
2311.294 Million cell updates/sec

Title: US-10-622-817-8
Perfect score: 2596
Sequence: 1 MIVFVRFNSSHGFPVEVDSD.....CWNCGCEWNRVCMGDHWFDV 465

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	490	18.9	357	2	T23460	hypothetical prote
2	220	8.5	543	2	H84724	probable ARI-like
3	219.5	8.5	1753	2	T00350	hypothetical prote
4	205	7.9	437	2	T16477	hypothetical prote
5	205	7.9	491	2	F87793	protein C27A12.6 [
6	198.5	7.6	644	2	T02366	hypothetical prote
7	196.5	7.6	514	2	G84724	probable ARI-like
8	195	7.5	320	2	T45909	hypothetical prote
9	193	7.4	542	2	A84725	similar to Ariadne
10	190.5	7.3	436	2	H87793	protein C27A12.8 [
11	188	7.2	1048	2	T31653	hypothetical prote
12	186.5	7.2	324	2	T16983	hypothetical prote
13	182.5	7.0	551	2	S38086	hypothetical prote
14	180.5	7.0	594	2	T04783	hypothetical prote
15	178.5	6.9	497	2	G87793	protein C27A12.7 [
16	178.5	6.9	688	2	B86448	hypothetical prote
17	174.5	6.7	408	2	T47498	hypothetical prote
18	174	6.7	348	2	T47494	hypothetical prote
19	173	6.7	565	2	F84721	probable RING zinc
20	162.5	6.3	498	2	JC5983	protein kinase C-i
21	159	6.1	518	2	E84536	hypothetical prote
22	158.5	6.1	451	2	T16481	hypothetical prote
23	157	6.0	488	2	T29562	hypothetical prote
24	155	6.0	532	2	T04748	hypothetical prote
25	152.5	5.9	816	2	T25555	hypothetical prote
26	149	5.7	464	2	S48329	probable membrane
27	147	5.7	1209	2	T52523	hypothetical prote
28	143.5	5.5	155	2	JH0227	ubiquitin / riboso
29	142.5	5.5	155	2	JH0226	ubiquitin / riboso
30	141.5	5.5	468	2	A84601	Mutator-like trans
31	136	5.2	222	2	T47487	hypothetical prote
32	134.5	5.2	229	1	UQJNI	ubiquitin precurso
33	134	5.2	373	2	E84647	hypothetical prote
34	134	5.2	638	2	T02611	hypothetical prote
35	133	5.1	397	2	H84578	probable RING zinc
36	132.5	5.1	869	1	JC4858	VLDL receptor prec
37	132	5.1	80	2	T28305	ORF MSV144 probabl
38	132	5.1	150	2	T30390	probable ubiquitin
39	132	5.1	154	2	S55242	polyubiquitin 2 -
40	132	5.1	156	2	C86439	protein T19E23.13
41	131	5.0	534	2	S34285	polyubiquitin - to
42	130.5	5.0	305	1	S29853	polyubiquitin 4 -
43	130.5	5.0	356	1	UQUTRC	polyubiquitin / ri
44	130	5.0	155	2	S40240	ubiquitin/ribosoma
45	130	5.0	156	2	T52335	ubiquitin extensio

ALIGNMENTS

RESULT 1
T23460
hypothetical protein K08E3.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T23460
R;McMurray, A.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19743
A;Accession: T23460
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-357 <WIL>
A;Cross-references: UNIPROT:Q9XUS3; EMBL:Z81568; PIDN:CAB04599.1; GSPDB:GN00021; CESP:K08
A;Experimental source: clone K08E3
C;Genetics:
A;Gene: CESP:K08E3.7
A;Map position: 3
A;Introns: 23/3; 72/2; 218/1; 286/2; 311/2

Query Match 18.9%; Score 490; DB 2; Length 357;
Best Local Similarity 26.7%; Pred. No. 1.8e-29;
Matches 119; Conservative 53; Mismatches 154; Indels 120; Gaps 12;

```
Qy         22 SIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCDLDQQSIVHIVQRPWRKGQEMN 81
              :|  | : | |   :|:|:| |:| ||:|     :::  |  : : :::  :    |
Db         29 NIEDLTKDVEKLTEIPSDELEVVFCGKKLSKSTIMRDLSLTPATQIMLLRPKFNSHNENG 88

Qy         82 ATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLAVILHTDSRKDSPPAGSPAGRS 141
              ||                                    : |||             |
Db         89 AT----------------------------------TAKITTDS------------S 99

Qy        142 IYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGPSCWDDVLIPNRMSGECQSPHC 201
              |  ||||:||  |  |: |||||  |  |   :: :   |  | |||   |:   |:
Db        100 ILGSFYVWCKN-CDDVKRGKLRVYCQKCSSTSVLVKSEPQNWSDVLKSKRIPAVCEECCT 158

Qy        202 PGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCITCTDVRSPVLVFQCNSRHVICL 261
              ||  ||| |||     :  :   ||  :  | :   |  |      :    ||  |: |
Db        159 PGLFAEFKFKC----LACNDPAAALTHVRGNWQMTECCVCDGKEKVIFDLGCN--HITCQ 212

Qy        262 DCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKELHHFRILGEEQYNRYQQYGAE 321
              || :                                :::::||| |:|:  |: ||:   |
Db        213 FCFRV----------------------------RVVQDVHHFHIMGQTSYSEYQRKATE 243

Qy        322 ECV-LQMGGVLCPRPGCGAGLLPEP--DQRKVTCEGGNGLGCGFAFCRECKEAYHEGECS 378
                : :   || ||   ||     ||  |  :  |       | |:|||:|     |  |
Db        244 RLIAVDDKGVTCPNVSCGQSFFWEPYDDDGRSQCP-----DCFFSPCRKC----FERNCV 294

Qy        379 AVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPRCHVPVEKNGGCMHMKCPQ 438
                 |  | |                        ||  ||: ||:|||  |:|||| |: |
Db        295 CQSEDDLTRT--------------------TIDATTRRCPKCHVATERNGGCAHIHC-- 331
```

```
Qy        439 PQCRLEWCWNCGCEWNRVCMGDHWFD 464
Db        332 TSCGMDWCFKCKTEWKEECQWDHWFN 357
```

RESULT 2
H84724
probable ARI-like RING zinc finger protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: H84724
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L. euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84724
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-543 <STO>
A;Cross-references: UNIPROT:Q9SKC3; GB:AE002093; NID:g4887759; PIDN:AAD32295.1; GSPDB:GN
C;Genetics:
A;Gene: At2g31770
A;Map position: 2

Query Match 8.5%; Score 220; DB 2; Length 543;
Best Local Similarity 25.9%; Pred. No. 5.9e-09;
Matches 60; Conservative 33; Mismatches 83; Indels 56; Gaps 11;

```
Qy        235 NITCITCTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLND-----RQFVHDPQLGYSLPC 289
Db        125 NIQCGICFESYTREEIARVSCGHPYCKTCWAGYITTKIEDGPGCLRVKCPEPS------C 178
Qy        290 VAGCPNSLIKELHHFRILGEEQYNRY--QQYGAEECVLQMGGVLCPRPGCGAGLL---PE 344
Db        179 SAAVGKDMIEDVTETKV--NEKYSRYILRSYVEDGKKIKW----CPSPGCGYAVEFGGSE 232
Qy        345 PDQRKVTCEGGNGLGCGFAFCRECKEAYHEG-ECSAVFEASGTTTQAYRVDERAAEQARW 403
Db        233 SSSYDVSCL------CSYRFCWNCSEDAHSPVDCDTV----------------------SKW 266
Qy        404 ------EAASKETIKKTTKPCPRCHVPVEKNGGCMHMKCPQPQCRLEWCWNC 449
Db        267 IFKNQDESENKNWMLANSKPCPECKRPIEKNDGCNHMTCSAP-CGHEFCWIC 317
```

RESULT 3
T00350
hypothetical protein KIAA0708 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C;Accession: T00350
R;Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N. DNA Res. 5, 169-176, 1998
A;Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A;Reference number: Z14142; MUID:98403880; PMID:9734811
A;Accession: T00350
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1753 <ISH>
A;Cross-references: UNIPROT:Q8IWT3; EMBL:AB014608; NID:g3327229; PIDN:BAA31683.1; PID:g3
A;Experimental source: brain
C;Genetics:
A;Note: KIAA0708

Query Match 8.5%; Score 219.5; DB 2; Length 1753;
Best Local Similarity 21.4%; Pred. No. 2.2e-08;
Matches 125; Conservative 57; Mismatches 181; Indels 221; Gaps 32;

```
Qy          1 MIVFVRFNSSHGFPVEV---DSDTS----------------------------IFQLK 27
Db       1012 MWLLLKFNQTEEVSVETLLKDSDLSPELLLQALVPLTSGNGPLTLHEGQDFPHGGVLRLH 1071
```

```
Qy         28 EVVAKRQG------------VPADQLRVIFAGKELRNDWTV--------QNCDLDQQSI 66
Db       1072 EPGPQRSGEALWLIPPQAYLNVEKDEGRTLEQKRNLLSCLLVRILKAHGEKGLHIDQ--L 1129
Qy         67 VHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLAVILHTD 126
Db       1130 VCLVLEAWQKGPNPPGTLG---HTVAGGV-----ACTSTDVLSCIL--HLLGQGYV---- 1175
Qy        127 SRKDSPP------AGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
Db       1176 KRRDDRPQILMYAAPEPMGPCRGQADVPFCGSQSETSKPSPEAV-------ATLASLQLP 1228
Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFK---CGAHPTSDKETPVALHLIATNS---- 233
Db       1229 -----AGRTMSPQ----EVEGLMKQTVRQVQETLNLEPDVAQHLLAHSHWGAE 1272
Qy        234 ----------------------------RNITCITCTDVRSPV-----LVFQCNSRHVI 259
Db       1273 QLLQSYSEDPEPLLLAAGLCVHQAQAVPVRPDHCPVCV---SPLGCDDDLPSLC-CMHYC 1328
Qy        260 CLDCFHLYCVTRLNDRQFVHDPQLGYSLPC-VAGCP----NSLIKELHHFRILGEEQYNR 314
Db       1329 CKSCWNEYLTTRI-------EQNLVLNCTCPIADCPAQPTGAFIRAI----VSSPEVISK 1377
Qy        315 YQQ-----YGAEECVLQMGGVLCPRP-GCGAGLLPEPDQRKVTCEGGNGLG-----CGFA 363
Db       1378 YEKALLRGY-VESC---SNLTWCTNPQGCD----------RILCRQGLGCGTTCSKCGWA 1423
Qy        364 FCRECK--EAYHEGECSAVFEASGTTTQAYRVDERAAEQARW-----------EAASKE 409
Db       1424 SCFNCSFPEAHYPASC-------------------GHMSQWVDDGGYYDGMSVEAQSKH 1463
Qy        410 TIKKTTKPCPRCHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEW 453
Db       1464 LAKLISKRCPSCQAPIEKNEGCLHMTC--AKCNHGFCWRCLKSW 1505
```

RESULT 4
T16477
hypothetical protein F56D2.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T16477
R;Du, Z.
submitted to the EMBL Data Library, August 1994
A;Description: The sequence of C. elegans cosmid F56D2.
A;Reference number: Z18519
A;Accession: T16477
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-437 <DUZ>
A;Cross-references: UNIPROT:Q20871; EMBL:U13644; NID:g532100; PID:g1945502; PIDN:AAB5268
A;Experimental source: strain Bristol N2; clone F56D2
C;Genetics:
A;Gene: CESP:F56D2.5
A;Map position: 3
A;Introns: 41/3; 134/3; 185/3; 223/3; 254/2; 287/2; 414/3

Query Match 7.9%; Score 205; DB 2; Length 437;
Best Local Similarity 20.2%; Pred. No. 6.3e-08;
Matches 104; Conservative 66; Mismatches 169; Indels 176; Gaps 28;

```
Qy         18 DSDTSIFQLKEVVAKRQGVPADQLRVIFAGKEL--RNDWTVQNCDLDQQSIVHIVQRPWR 75
Db          3 DRDLQIYEL-----------EALESVLREKKLAKSSDWSDKNAEI--QGIIEV------ 42
Qy         76 KGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLAVILHTDSRKDSPPAG 135
Db         43 --------GFDNLYDPTVTIEGTSDSGDQFHLPLDILP------PIRLKFHLPNDYPTVS 88
Qy        136 SPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGPSCWDDVLIPNRMSGE 195
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:25:39 ; Search time 102.258 Seconds
(without alignments)
1758.725 Million cell updates/sec

Title: US-10-622-817-9
Perfect score: 2596
Sequence: 1 MIVFVRFNSSHGFPVEVDSD.....CWNCGCEWNRVCMGDHWFDV 465

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2596	100.0	465	2	AAY32501	Aay32501 Human par
2	2596	100.0	465	6	AAE30800	Aae30800 Human Par
3	2596	100.0	465	6	ABO07156	Abo07156 Human p53
4	2584	99.5	465	6	AAE30801	Aae30801 Human Par
5	2405	92.6	437	2	AAY32502	Aay32502 Human par
6	2405	92.6	437	6	ABO07157	Abo07157 Human p53
7	2170.5	83.6	464	4	AAB67517	Aab67517 Amino aci
8	2165.5	83.4	464	4	AAB67531	Aab67531 Amino aci
9	2165.5	83.4	464	4	AAB67532	Aab67532 Amino aci
10	2087.5	80.4	451	4	AAB67533	Aab67533 Amino aci
11	1696.5	65.4	316	6	ABO07158	Abo07158 Human p53
12	1679.5	64.7	344	4	AAB67521	Aab67521 Amino aci
13	1234.5	47.6	296	4	AAB67526	Aab67526 Amino aci
14	1058.5	40.8	468	4	ABB65114	Abb65114 Drosophil
15	1002	38.6	250	4	AAB67519	Aab67519 Amino aci
16	997	38.4	262	4	AAB67518	Aab67518 Amino aci
17	870	33.5	156	4	AAG67213	Aag67213 Amino aci
18	671.5	25.9	183	4	AAB67525	Aab67525 Amino aci
19	671.5	25.9	194	4	AAB67524	Aab67524 Amino aci
20	544.5	21.0	386	4	AAB86951	Aab86951 C. elegan
21	506	19.5	153	4	AAB67523	Aab67523 Amino aci
22	335	12.9	77	4	AAB67529	Aab67529 Amino aci
23	263	10.1	46	4	AAM17055	Aam17055 Peptide #
24	263	10.1	46	4	ABB36052	Abb36052 Peptide #
25	263	10.1	46	4	AAM29548	Aam29548 Peptide #
26	263	10.1	46	4	ABB30873	Abb30873 Peptide #
27	263	10.1	46	4	ABB21452	Abb21452 Protein #
28	263	10.1	46	4	AAM69222	Aam69222 Human bon
29	263	10.1	46	4	AAM56839	Aam56839 Human bra
30	263	10.1	46	4	ABG50898	Abg50898 Human liv
31	263	10.1	46	5	ABG38830	Abg38830 Human pep
32	252	9.7	106	8	ADM96489	Adm96489 Modified
33	251	9.7	63	4	AAB67522	Aab67522 Amino aci
34	251	9.7	105	4	AAB67520	Aab67520 Amino aci
35	251	9.7	105	8	ADM96487	Adm96487 Modified
36	240.5	9.3	503	4	ABB61708	Abb61708 Drosophil
37	240.5	9.3	503	7	ADB67082	Adb67082 Ariadne-1
38	240	9.2	511	4	ABB63665	Abb63665 Drosophil
39	239	9.2	520	8	ADP22562	Adp22562 Sea-squir
40	236.5	9.1	445	3	AAB21034	Aab21034 Human nuc
41	236.5	9.1	445	8	ADQ96154	Adq96154 T cell ac
42	236.5	9.1	445	8	ADQ96026	Adq96026 T cell ac
43	236.5	9.1	557	3	AAY98059	Aay98059 Human Rin
44	236.5	9.1	557	8	ADQ96028	Adq96028 T cell ac
45	236	9.1	53	4	AAB67528	Aab67528 Amino aci

ALIGNMENTS

RESULT 1
AAY32501
ID AAY32501 standard; protein; 465 AA.
XX
AC AAY32501;
XX
DT 21-OCT-1999 (first entry)
XX
DE Human parkin gene variant protein.
XX
KW Parkinson's disease related gene; parkin gene; variant; gene therapy.
XX
OS Homo sapiens.
XX
PN WO9940191-A1.
XX
PD 12-AUG-1999.
XX
PF 09-FEB-1999; 99WO-JP000545.
XX
PR 09-FEB-1998; 98JP-00027531.
XX
PA (SHIM/) SHIMIZU N.
PA (MIZU/) MIZUNO Y.
XX
PI Shimizu N, Mizuno Y;
XX
DR WPI; 1999-494295/41.
DR N-PSDB; AAX99923.
XX
PT Gene implicated in the pathology of Parkinson's disease, used for
PT treatment of the disease.
XX
PS Claim 1; Page 83-88; 114pp; English.
XX
CC This sequence is encoded by a gene of the invention, and is implicated in
CC the pathology of Parkinson's disease. This sequence is a variant of the
CC parkin gene found in parkinson's disease patients. The sequences may be
CC used for the diagnosis, treatment (including gene therapy) and
CC investigation of Parkinson's disease
XX
SQ Sequence 465 AA;

Query Match 100.0%; Score 2596; DB 2; Length 465;
Best Local Similarity 100.0%; Pred. No. 2.3e-219;
Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

```
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180

Db        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240

Db        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300

Db        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360

Db        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420

Db        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465

Db        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
```

RESULT 2
AAE30800
ID AAE30800 standard; protein; 465 AA.
XX
AC AAE30800;
XX
DT 24-FEB-2003 (first entry)
XX
DE Human Parkin protein.
XX
KW Human; Parkin protein; neurological disorder; apoptosis; gene therapy;
KW ischaemic stroke; Parkinson's disease; Alzheimer's disease; nootropic;
KW transgenic; cerebroprotective; neuroprotective; neurotransplantation.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Cleavage-site 126
XX
PN WO200279459-A2.
XX
PD 10-OCT-2002.
XX
PF 02-APR-2002; 2002WO-DK000221.
XX
PR 29-MAR-2001; 2001DK-00000525.
PR 03-APR-2001; 2001US-0281286P.
XX
PA (NSGE-) NSGENE AS.
XX
PI Jensen PH;
XX
DR WPI; 2003-046812/04.
DR N-PSDB; AAD47679.
XX
PT New isolated nucleic acid sequence encoding a Parkin polypeptide, useful
PT for treating, preventing or diagnosing neurological disorders, e.g.
PT Parkinson's disease, Alzheimer's disease or ischemic stroke, and in
PT screening assays.
XX
PS Claim 10; Page 69; 71pp; English.
XX
CC The invention relates to Parkin protein and its corresponding nucleic
CC acid sequence. The nucleic acid sequence is useful for altering the
CC proteolytic processing of Parkin at its potential cleavage site at Asp
CC 126. The invention is used in manufacturing or testing a pharmaceutical
CC composition for treating and/or preventing a neurological disorder, e.g.
CC Alzheimer's disease or ischaemic stroke. It also used for detecting the
CC occurrence of proteolytic processing of Parkin at Asp 126 in a sample, in
CC monitoring a potential disposition for a neurodegenerative disease, and
CC for treating, preventing and/or diagnosing Parkinson's disease or other
CC neurodegenerative disorders. The viral vector is used for transforming
CC neuronal cells in vivo or ex vivo. The invention is useful for
CC neurotransplantation into the CNS of a mammal. It may be used in
CC screening assays to identify compounds that increase or decrease
CC apoptosis. It is also used in gene therapy. The present sequence is human
CC Parkin protein
XX
SQ Sequence 465 AA;

Query Match 100.0%; Score 2596; DB 6; Length 465;
Best Local Similarity 100.0%; Pred. No. 2.3e-219;
Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180

Db        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240

Db        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300

Db        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360

Db        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420

Db        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465

Db        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
```

RESULT 3
ABO07156
ID ABO07156 standard; protein; 465 AA.
XX
AC ABO07156;
XX
DT 13-AUG-2003 (first entry)
XX
DE Human p53 modifying protein, SEQ ID 116.
XX
KW Human; p53 modifier; cytostatic; cancer; cytostatic; antiangiogenic;
KW antiapoptotic; p53 pathway; breast cancer; colon cancer; kidney cancer;
KW lung cancer; ovarian cancer; angiogenesis; cell cycle;
KW apoptotic disorder; cell proliferation disorder.
XX
OS Homo sapiens.
XX

PN WO200299122-A1.
XX
PD 12-DEC-2002.
XX
PF 03-JUN-2002; 2002WO-US017382.
XX
PR 05-JUN-2001; 2001US-0296076P.
PR 10-OCT-2001; 2001US-0328605P.
PR 15-FEB-2002; 2002US-0357253P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX
DR WPI; 2003-156859/15.
DR N-PSDB; ACD13332.
XX
PT Identifying modulators of the p53 pathway for use in treating apoptotic
PT or cell proliferation disorders, comprises screening for agents that
PT modulate activity of a human ortholog of genes that modify the p53
PT pathway in Drosophila.
XX
PS Example 2; Page 399-401; 678pp; English.
XX
CC The invention relates to identifying (M1) a candidate p53 pathway
CC modulating agent, by contacting an assay system comprising a purified HM
CC polypeptide (human orthologue of genes that modify the p53 pathway in
CC Drosophila) or nucleic acid with a test agent under conditions, where but
CC for the presence of the test agent, the system provides a reference
CC activity, and detecting a test agent-biased activity of the assay system.
CC Also included are modulating (M2) a p53 pathway of a cell (comprising
CC contacting a cell defective in p53 function with a candidate modulator
CC that specifically binds to a HM polypeptide comprising an HM amino acid
CC sequence, where p53 function is restored), modulating (M3) a p53 pathway
CC in a mammalian cell (comprising contacting the cell with an agent that
CC specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)
CC a disease in a patient (comprising: (a) obtaining a biological sample
CC from the patient; (b) contacting the sample with a probe for HM
CC expression; (c) comparing the results with a control; and (d) determining
CC whether the comparison indicates a likelihood disease). (M1) is useful
CC for identifying modulators of the p53 pathway. A probe for HM expression
CC is useful for diagnosing breast, colon, kidney, lung and ovarian cancer,
CC in a patient, where the cancer has greater than 25 % expression level.
CC Modulators identified by (M1) are useful in a variety of diagnostic and
CC therapeutic applications, where disease or disorder prognosis is related
CC to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell
CC proliferation disorders (e.g. cancer). Another two new methods (M2 and
CC M3) are useful for modulating the p53 pathway of a cell, thus restoring
CC the p53 function of the cell, so that the cell undergoes normal
CC proliferation or progression through the cell cycle. (M2) and (M3) are
CC also useful for treating defects in the p53 pathway such as angiogenic,
CC apoptotic or cell proliferation disorders. The present sequence
CC represents a human p53 pathway modifying protein
XX
SQ Sequence 465 AA;

Query Match 100.0%; Score 2596; DB 6; Length 465;
Best Local Similarity 100.0%; Pred. No. 2.3e-219;
Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |||||||||||||||||||||||||||||||||||||||||||||
Db        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
```

RESULT 4
AAE30801
ID AAE30801 standard; protein; 465 AA.
XX
AC AAE30801;
XX
DT 24-FEB-2003 (first entry)
XX
DE Human Parkin D126 mutant protein.
XX
KW Human; Parkin protein; neurological disorder; apoptosis; gene therapy;
KW ischaemic stroke; Parkinson's disease; Alzheimer's disease; nootropic;
KW transgenic; cerebroprotective; neuroprotective; neurotransplantation;
KW mutant; mutein.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 126
FT /note= "Wild-type Asp is replaced with Glu"
FT Cleavage-site 126
FT Misc-difference 223
FT /note= "Encoded by CCA"
XX
PN WO200279459-A2.
XX
PD 10-OCT-2002.
XX
PF 02-APR-2002; 2002WO-DK000221.
XX
PR 29-MAR-2001; 2001DK-00000525.
PR 03-APR-2001; 2001US-0281286P.
XX
PA (NSGE-) NSGENE AS.
XX
PI Jensen PH;
XX
DR WPI; 2003-046812/04.
DR N-PSDB; AAD47680.
XX
PT New isolated nucleic acid sequence encoding a Parkin polypeptide, useful
PT for treating, preventing or diagnosing neurological disorders, e.g.
PT Parkinson's disease, Alzheimer's disease or ischemic stroke, and in
PT screening assays.
XX
PS Claim 10; Page 71; 71pp; English.
XX
CC The invention relates to Parkin protein and its corresponding nucleic
CC acid sequence. The nucleic acid sequence is useful for altering the
CC proteolytic processing of Parkin at its potential cleavage site at Asp
CC 126. The invention is used in manufacturing or testing a pharmaceutical
CC composition for treating and/or preventing a neurological disorder, e.g.

CC Alzheimer's disease or ischaemic stroke. It also used for detecting the
CC occurrence of proteolytic processing of Parkin at Asp 126 in a sample, in
CC monitoring a potential disposition for a neurodegenerative disease, and
CC for treating, preventing and/or diagnosing Parkinson's disease or other
CC neurodegenerative disorders. The viral vector is used for transforming
CC neuronal cells in vivo or ex vivo. The invention is useful for
CC neurotransplantation into the CNS of a mammal. It may be used in
CC screening assays to identify compounds that increase or decrease
CC apoptosis. It is also used in gene therapy. The present sequence is human
CC Parkin mutant protein
XX
SQ Sequence 465 AA;

Query Match 99.5%; Score 2584; DB 6; Length 465;
Best Local Similarity 99.6%; Pred. No. 2.6e-218;
Matches 463; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VILHTESRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETSVALHLIATNSRNITCIT 240

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |||||||||||||||||||||||||||||||||||||||||||||
Db        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
```

RESULT 5
AAY32502
ID AAY32502 standard; protein; 437 AA.
XX
AC AAY32502;
XX
DT 21-OCT-1999 (first entry)
XX
DE Human parkin gene variant protein.
XX
KW Parkinson's disease related gene; parkin gene; variant; gene therapy.
XX
OS Homo sapiens.
XX
PN WO9940191-A1.
XX
PD 12-AUG-1999.
XX
PF 09-FEB-1999; 99WO-JP000545.
XX
PR 09-FEB-1998; 98JP-00027531.
XX
PA (SHIM/) SHIMIZU N.
PA (MIZU/) MIZUNO Y.
XX
PI Shimizu N, Mizuno Y;
XX
DR WPI; 1999-494295/41.
DR N-PSDB; AAX99924.
XX
PT Gene implicated in the pathology of Parkinson's disease, used for
PT treatment of the disease.
XX
PS Claim 1; Page 89-94; 114pp; English.
XX
CC This sequence is encoded by a gene of the invention, and is implicated in
CC the pathology of Parkinson's disease. This sequence is a variant of the
CC parkin gene found in parkinson's disease patients. The sequences may be
CC used for the diagnosis, treatment (including gene therapy) and
CC investigation of Parkinson's disease
XX
SQ Sequence 437 AA;

Query Match 92.6%; Score 2405; DB 2; Length 437;
Best Local Similarity 94.0%; Pred. No. 1.3e-202;
Matches 437; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

```
Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQ-- 178

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
                                          ||||||||||||||||||||||||||||||||
Db        179 --------------------------EFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 212

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 272

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 332

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 392

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |||||||||||||||||||||||||||||||||||||||||||||
Db        393 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 437
```

RESULT 6
ABO07157
ID ABO07157 standard; protein; 437 AA.
XX
AC ABO07157;
XX
DT 13-AUG-2003 (first entry)
XX
DE Human p53 modifying protein, SEQ ID 117.
XX
KW Human; p53 modifier; cytostatic; cancer; cytostatic; antiangiogenic;
KW antiapoptotic; p53 pathway; breast cancer; colon cancer; kidney cancer;
KW lung cancer; ovarian cancer; angiogenesis; cell cycle;
KW apoptotic disorder; cell proliferation disorder.
XX
OS Homo sapiens.

XX
PN WO200299122-A1.
XX
PD 12-DEC-2002.
XX
PF 03-JUN-2002; 2002WO-US017382.
XX
PR 05-JUN-2001; 2001US-0296076P.
PR 10-OCT-2001; 2001US-0328605P.
PR 15-FEB-2002; 2002US-0357253P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX
DR WPI; 2003-156859/15.
DR N-PSDB; ACD13333.
XX
PT Identifying modulators of the p53 pathway for use in treating apoptotic
PT or cell proliferation disorders, comprises screening for agents that
PT modulate activity of a human ortholog of genes that modify the p53
PT pathway in Drosophila.
XX
PS Example 2; Page 401-402; 678pp; English.
XX
CC The invention relates to identifying (M1) a candidate p53 pathway
CC modulating agent, by contacting an assay system comprising a purified HM
CC polypeptide (human orthologue of genes that modify the p53 pathway in
CC Drosophila) or nucleic acid with a test agent under conditions, where but
CC for the presence of the test agent, the system provides a reference
CC activity, and detecting a test agent-biased activity of the assay system.
CC Also included are modulating (M2) a p53 pathway of a cell (comprising
CC contacting a cell defective in p53 function with a candidate modulator
CC that specifically binds to a HM polypeptide comprising an HM amino acid
CC sequence, where p53 function is restored), modulating (M3) a p53 pathway
CC in a mammalian cell (comprising contacting the cell with an agent that
CC specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)
CC a disease in a patient (comprising: (a) obtaining a biological sample
CC from the patient; (b) contacting the sample with a probe for HM
CC expression; (c) comparing the results with a control; and (d) determining
CC whether the comparison indicates a likelihood disease). (M1) is useful
CC for identifying modulators of the p53 pathway. A probe for HM expression
CC is useful for diagnosing breast, colon, kidney, lung and ovarian cancer,
CC in a patient, where the cancer has greater than 25 % expression level.
CC Modulators identified by (M1) are useful in a variety of diagnostic and
CC therapeutic applications, where disease or disorder prognosis is related
CC to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell
CC proliferation disorders (e.g. cancer). Another two new methods (M2 and
CC M3) are useful for modulating the p53 pathway of a cell, thus restoring
CC the p53 function of the cell, so that the cell undergoes normal
CC proliferation or progression through the cell cycle. (M2) and (M3) are
CC also useful for treating defects in the p53 pathway such as angiogenic,
CC apoptotic or cell proliferation disorders. The present sequence
CC represents a human p53 pathway modifying protein
XX
SQ Sequence 437 AA;

Query Match 92.6%; Score 2405; DB 6; Length 437;
Best Local Similarity 94.0%; Pred. No. 1.3e-202;
Matches 437; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

```
Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQ-- 178

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
                                            ||||||||||||||||||||||||||||||
Db        179 -------------------------EFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 212

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 272

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 332

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 392

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |||||||||||||||||||||||||||||||||||||||||||||
Db        393 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 437
```

RESULT 7
AAB67517
ID AAB67517 standard; protein; 464 AA.
XX
AC AAB67517;
XX
DT 29-MAY-2001 (first entry)
XX
DE Amino acid sequence of a murine parkin2 polypeptide.
XX
KW Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW brain tumour; head trauma; stroke; vascular irregularity;
KW metabolic irregularity.
XX
OS Mus sp.
XX
PN EP1081225-A1.
XX
PD 07-MAR-2001.
XX
PF 30-AUG-1999; 99EP-00116766.
XX
PR 30-AUG-1999; 99EP-00116766.
XX
PA (BIOF-) BIOFRONTERA PHARM GMBH.
XX
PI Luebbert H;
XX
DR WPI; 2001-212797/22.
DR N-PSDB; AAF55244.
XX
PT New polynucleotides encoding mouse parkin2 protein, useful for producing
PT a transgenic non-human animal as an animal model for neurodegenerative
PT diseases.
XX
PS Disclosure; Page 17-19; 62pp; English.
XX
CC The present sequence represents a murine parkin2 polypeptide. Mutations
CC or deletions in the parkin2 gene cause Parkinson's disease in humans. The
CC human parkin2 gene is located in gene region 6q25.2-27. Parkin2
CC polypeptides and polynucleotides are useful for analysing
CC neurodegenerative diseases. They are also useful for testing the efficacy
CC of the treatment of a neurodegenerative disease such as Parkinson's
CC disease, Alzheimer's disease, Huntington's disease, amyotrophic lateral
CC sclerosis, Multi-system atrophy, Wilson's disease, Pick's disease, Prion
CC disease, and secondary causes inducing Parkinson's syndromes like toxins,
CC drugs, brain tumours, head trauma, stroke, vascular irregularities or
CC metabolic irregularities, associated with a less active or non-active
CC parkin protein
XX

```
SQ   Sequence 464 AA;

  Query Match             83.6%;  Score 2170.5;  DB 4;  Length 464;
  Best Local Similarity   83.4%;  Pred. No. 6e-182;
  Matches  388;  Conservative   29;  Mismatches   47;  Indels    1;  Gaps    1;

Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||:|||||||||||| ||||||||||||||||||||||||| |  ||||||
Db          1 MIVFVRFNSSYGFPVEVDSDTSILQLKEVVAKRQGVPADQLRVIFAGKELPNHLTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              |:||||||||||| |:  | ||:|||:|:: : |   | :|||||||||  || ||||||
Db         61 LEQQSIVHIVQRPRRRSHETNASGGDEPQSTSEGSIWESRSLTRVDLSSHTLPVDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||| |||::||  |  |  :  ||||::|||||| :|||||||||| ||:|||||| |||
Db        121 VILDTDSKRDSEAARGPV-KPTYNSFFIYCKGPCHKVQPGKLRVQCGTCKQATLTLAQGP 179

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              |||||||||||||||||| |||| |||||||||||||:| |||:|| :| |:| ||
Db        180 SCWDDVLIPNRMSGECQSPDCPGTRAEFFFKCGAHPTSDKDTSVALNLITSNRRSIPCIA 239

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              |||||||||||||| |||||||||||||||||||||||||| ||||||||||||||||||
Db        240 CTDVRSPVLVFQCNHRHVICLDCFHLYCVTRLNDRQFVHDAQLGYSLPCVAGCPNSLIKE 299

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              |||||||||||| |||||||||||||||||||||||||||||||| ||||||||||||||
Db        300 LHHFRILGEEQYTRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEQGQRKVTCEGGNGLGC 359

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              || |||:||||||||:| :: | || |:||||||:||||||||| ||||||||||||||
Db        360 GFVFCRDCKEAYHEGDCDSLLEPSGATSQAYRVDKRAAEQARWEEASKETIKKTTKPCPR 419

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |:||:|||||||||||||||:|||||||||||| |||||||||
Db        420 CNVPIEKNGGCMHMKCPQPQCKLEWCWNCGCEWNRACMGDHWFDV 464


RESULT 8
AAB67531
ID   AAB67531 standard; protein; 464 AA.
XX
AC   AAB67531;
XX
DT   29-MAY-2001  (first entry)
XX
DE   Amino acid sequence of a mutated murine parkin2 polypeptide.
XX
KW   Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW   Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW   Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW   brain tumour; head trauma; stroke; vascular irregularity;
KW   metabolic irregularity.
XX
OS   Mus sp.
XX
PN   EP1081225-A1.
XX
PD   07-MAR-2001.
XX
PF   30-AUG-1999;   99EP-00116766.
XX
PR   30-AUG-1999;   99EP-00116766.
XX
PA   (BIOF-) BIOFRONTERA PHARM GMBH.
XX
PI   Luebbert H;
XX
DR   WPI; 2001-212797/22.
DR   N-PSDB; AAF55258.
XX
PT   New polynucleotides encoding mouse parkin2 protein, useful for producing
PT   a transgenic non-human animal as an animal model for neurodegenerative
PT   diseases.
XX
PS   Claim 7; Page 47-49; 62pp; English.
XX
CC   The present sequence represents a murine parkin2 polypeptide. The
CC   sequence contains the mutation Lys161Asn. Mutations or deletions in the
CC   parkin2 gene cause Parkinson's disease in humans. The human parkin2 gene
CC   is located in gene region 6q25.2-27. Parkin2 polypeptides and
CC   polynucleotides are useful for analysing neurodegenerative diseases. They
CC   are also useful for testing the efficacy of the treatment of a
CC   neurodegenerative disease such as Parkinson's disease, Alzheimer's
CC   disease, Huntington's disease, amyotrophic lateral sclerosis, Multi-
CC   system atrophy, Wilson's disease, Pick's disease, Prion disease, and
CC   secondary causes inducing Parkinson's syndromes like toxins, drugs, brain
CC   tumours, head trauma, stroke, vascular irregularities or metabolic
CC   irregularities, associated with a less active or non-active parkin
CC   protein
XX
SQ   Sequence 464 AA;

  Query Match             83.4%;  Score 2165.5;  DB 4;  Length 464;
  Best Local Similarity   83.2%;  Pred. No. 1.7e-181;
  Matches  387;  Conservative   29;  Mismatches   48;  Indels    1;  Gaps    1;

Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||:|||||||||||| ||||||||||||||||||||||||| |  ||||||
Db          1 MIVFVRFNSSYGFPVEVDSDTSILQLKEVVAKRQGVPADQLRVIFAGKELPNHLTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              |:||||||||||| |:  | ||:|||:|:: : |   | :|||||||||  || ||||||
Db         61 LEQQSIVHIVQRPRRRSHETNASGGDEPQSTSEGSIWESRSLTRVDLSSHTLPVDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||| |||::||  |  |  :  ||||::|||||| :|||| ||||| ||:|||||| |||
Db        121 VILDTDSKRDSEAARGPV-KPTYNSFFIYCKGPCHKVQPGNLRVQCGTCKQATLTLAQGP 179

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              |||||||||||||||||| |||| |||||||||||||:| |||:|| :| |:| ||
Db        180 SCWDDVLIPNRMSGECQSPDCPGTRAEFFFKCGAHPTSDKDTSVALNLITSNRRSIPCIA 239

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              |||||||||||||| |||||||||||||||||||||||||| ||||||||||||||||||
Db        240 CTDVRSPVLVFQCNHRHVICLDCFHLYCVTRLNDRQFVHDAQLGYSLPCVAGCPNSLIKE 299

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              |||||||||||| |||||||||||||||||||||||||||||||| ||||||||||||||
Db        300 LHHFRILGEEQYTRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEQGQRKVTCEGGNGLGC 359

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              || |||:||||||||:| :: | || |:||||||:||||||||| ||||||||||||||
Db        360 GFVFCRDCKEAYHEGDCDSLLEPSGATSQAYRVDKRAAEQARWEEASKETIKKTTKPCPR 419

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |:||:|||||||||||||||:|||||||||||| |||||||||
Db        420 CNVPIEKNGGCMHMKCPQPQCKLEWCWNCGCEWNRACMGDHWFDV 464


RESULT 9
AAB67532
ID   AAB67532 standard; protein; 464 AA.
XX
AC   AAB67532;
XX
DT   29-MAY-2001  (first entry)
XX
DE   Amino acid sequence of a mutated murine parkin2 polypeptide.
XX
KW   Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW   Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW   Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW   brain tumour; head trauma; stroke; vascular irregularity;
KW   metabolic irregularity.
```

```
XX
OS   Mus sp.
XX
PN   EP1081225-A1.
XX
PD   07-MAR-2001.
XX
PF   30-AUG-1999;   99EP-00116766.
XX
PR   30-AUG-1999;   99EP-00116766.
XX
PA   (BIOF-) BIOFRONTERA PHARM GMBH.
XX
PI   Luebbert H;
XX
DR   WPI; 2001-212797/22.
DR   N-PSDB; AAF55259.
XX
PT   New polynucleotides encoding mouse parkin2 protein, useful for producing
PT   a transgenic non-human animal as an animal model for neurodegenerative
PT   diseases.
XX
PS   Claim 7; Page 49-51; 62pp; English.
XX
CC   The present sequence represents a murine parkin2 polypeptide. The
CC   sequence contains the mutation Thr415Asn. Mutations or deletions in the
CC   parkin2 gene cause Parkinson's disease in humans. The human parkin2 gene
CC   is located in gene region 6q25.2-27. Parkin2 polypeptides and
CC   polynucleotides are useful for analysing neurodegenerative diseases. They
CC   are also useful for testing the efficacy of the treatment of a
CC   neurodegenerative disease such as Parkinson's disease, Alzheimer's
CC   disease, Huntington's disease, amyotrophic lateral sclerosis, Multi-
CC   system atrophy, Wilson's disease, Pick's disease, Prion disease, and
CC   secondary causes inducing Parkinson's syndromes like toxins, drugs, brain
CC   tumours, head trauma, stroke, vascular irregularities or metabolic
CC   irregularities, associated with a less active or non-active parkin
CC   protein
XX
SQ   Sequence 464 AA;

  Query Match             83.4%;  Score 2165.5;  DB 4;  Length 464;
  Best Local Similarity   83.2%;  Pred. No. 1.7e-181;
  Matches  387;  Conservative   29;  Mismatches   48;  Indels    1;  Gaps    1;

Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||:|||||||||||||| |||||||||||||||||||||||| |  ||||||
Db          1 MIVFVRFNSSYGFPVEVDSDTSILQLKEVVAKRQGVPADQLRVIFAGKELPNHLTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              |:||||||||||| |:  | ||:|||:|:: : |   | :|||||||||  || ||||||
Db         61 LEQQSIVHIVQRPRRRSHETNASGGDEPQSTSEGSIWESRSLTRVDLSSHTLPVDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||| |||::||  |  |  :  ||||::|||||| :|||||||||| ||:|||||| |||
Db        121 VILDTDSKRDSEAARGPV-KPTYNSFFIYCKGPCHKVQPGKLRVQCGTCKQATLTLAQGP 179

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              |||||||||||||||||| |||| ||||||||||||||:| |||:|| :| |:| ||
Db        180 SCWDDVLIPNRMSGECQSPDCPGTRAEFFFKCGAHPTSDKDTSVALNLITSNRRSIPCIA 239

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              |||||||||||||| |||||||||||||||||||||||| ||||||||||||||||||||
Db        240 CTDVRSPVLVFQCNHRHVICLDCFHLYCVTRLNDRQFVHDAQLGYSLPCVAGCPNSLIKE 299

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              |||||||||||| |||||||||||||||||||||||||||||   ||||||||||||||
Db        300 LHHFRILGEEQYTRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEQGQRKVTCEGGNGLGC 359

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              || |||:|||||||:| :: | || |:||||||:||||||||| |||||||| |||||
Db        360 GFVFCRDCKEAYHEGDCDSLLEPSGATSQAYRVDKRAAEQARWEEASKETIKKTNKPCPR 419

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
```

```
              |:||:||||||||||||||:|||||||||||| |||||||||
Db        420 CNVPIEKNGGCMHMKCPQPQCKLEWCWNCGCEWNRACMGDHWFDV 464


RESULT 10
AAB67533
ID   AAB67533 standard; protein; 451 AA.
XX
AC   AAB67533;
XX
DT   29-MAY-2001  (first entry)
XX
DE   Amino acid sequence of a mutated murine parkin2 polypeptide.
XX
KW   Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW   Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW   Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW   brain tumour; head trauma; stroke; vascular irregularity;
KW   metabolic irregularity.
XX
OS   Mus sp.
XX
PN   EP1081225-A1.
XX
PD   07-MAR-2001.
XX
PF   30-AUG-1999;   99EP-00116766.
XX
PR   30-AUG-1999;   99EP-00116766.
XX
PA   (BIOF-) BIOFRONTERA PHARM GMBH.
XX
PI   Luebbert H;
XX
DR   WPI; 2001-212797/22.
DR   N-PSDB; AAF55260.
XX
PT   New polynucleotides encoding mouse parkin2 protein, useful for producing
PT   a transgenic non-human animal as an animal model for neurodegenerative
PT   diseases.
XX
PS   Claim 7; Page 51-53; 62pp; English.
XX
CC   The present sequence represents a murine parkin2 polypeptide. The
CC   sequence contains the mutation Trp453Stop. Mutations or deletions in the
CC   parkin2 gene cause Parkinson's disease in humans. The human parkin2 gene
CC   is located in gene region 6q25.2-27. Parkin2 polypeptides and
CC   polynucleotides are useful for analysing neurodegenerative diseases. They
CC   are also useful for testing the efficacy of the treatment of a
CC   neurodegenerative disease such as Parkinson's disease, Alzheimer's
CC   disease, Huntington's disease, amyotrophic lateral sclerosis, Multi-
CC   system atrophy, Wilson's disease, Pick's disease, Prion disease, and
CC   secondary causes inducing Parkinson's syndromes like toxins, drugs, brain
CC   tumours, head trauma, stroke, vascular irregularities or metabolic
CC   irregularities, associated with a less active or non-active parkin
CC   protein
XX
SQ   Sequence 451 AA;

  Query Match             80.4%;  Score 2087.5;  DB 4;  Length 451;
  Best Local Similarity   83.2%;  Pred. No. 1.2e-174;
  Matches  376;  Conservative   29;  Mismatches   46;  Indels    1;  Gaps    1;

Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||:|||||||||||||| |||||||||||||||||||||||| |  ||||||
Db          1 MIVFVRFNSSYGFPVEVDSDTSILQLKEVVAKRQGVPADQLRVIFAGKELPNHLTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              |:||||||||||| |:  | ||:|||:|:: : |   | :|||||||||  || ||||||
Db         61 LEQQSIVHIVQRPRRRSHETNASGGDEPQSTSEGSIWESRSLTRVDLSSHTLPVDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||| |||::||  |  |  :  ||||::|||||| :|||||||||| ||:|||||| |||
```

```
Db        121 VILDTDSKRDSEAARGPV-KPTYNSFFIYCKGPCHKVQPGKLRVQCGTCKQATLTLAQGP 179

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240

Db        180 SCWDDVLIPNRMSGECQSPDCPGTRAEFFFKCGAHPTSDKDTSVALNLITSNRRSIPCIA 239

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300

Db        240 CTDVRSPVLVFQCNHRHVICLDCFHLYCVTRLNDRQFVHDAQLGYSLPCVAGCPNSLIKE 299

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360

Db        300 LHHFRILGEEQYTRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEQGQRKVTCEGGNGLGC 359

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420

Db        360 GFVFCRDCKEAYHEGDCDSLLEPSGATSQAYRVDKRAAEQARWEEASKETIKKTTKPCPR 419

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCE 452

Db        420 CNVPIEKNGGCMHMKCPQPQCKLEWCWNCGCE 451
```

```
RESULT 11
ABO07158
ID   ABO07158 standard; protein; 316 AA.
XX
AC   ABO07158;
XX
DT   13-AUG-2003  (first entry)
XX
DE   Human p53 modifying protein, SEQ ID 118.
XX
KW   Human; p53 modifier; cytostatic; cancer; cytostatic; antiangiogenic;
KW   antiapoptotic; p53 pathway; breast cancer; colon cancer; kidney cancer;
KW   lung cancer; ovarian cancer; angiogenesis; cell cycle;
KW   apoptotic disorder; cell proliferation disorder.
XX
OS   Homo sapiens.
XX
PN   WO200299122-A1.
XX
PD   12-DEC-2002.
XX
PF   03-JUN-2002; 2002WO-US017382.
XX
PR   05-JUN-2001; 2001US-0296076P.
PR   10-OCT-2001; 2001US-0328605P.
PR   15-FEB-2002; 2002US-0357253P.
XX
PA   (EXEL-) EXELIXIS INC.
XX
PI   Friedman L,  Plowman GD,  Belvin M,  Francis-Lang H,  Li D,  Funke RP;
XX
DR   WPI; 2003-156859/15.
DR   N-PSDB; ACD13334.
XX
PT   Identifying modulators of the p53 pathway for use in treating apoptotic
PT   or cell proliferation disorders, comprises screening for agents that
PT   modulate activity of a human ortholog of genes that modify the p53
PT   pathway in Drosophila.
XX
PS   Example 2; Page 402-403; 678pp; English.
XX
CC   The invention relates to identifying (M1) a candidate p53 pathway
CC   modulating agent, by contacting an assay system comprising a purified HM
CC   polypeptide (human orthologue of genes that modify the p53 pathway in
CC   Drosophila) or nucleic acid with a test agent under conditions, where but
CC   for the presence of the test agent, the system provides a reference
CC   activity, and detecting a test agent-biased activity of the assay system.
CC   Also included are modulating (M2) a p53 pathway of a cell (comprising
CC   contacting a cell defective in p53 function with a candidate modulator
CC   that specifically binds to a HM polypeptide comprising an HM amino acid
CC   sequence, where p53 function is restored), modulating (M3) a p53 pathway
CC   in a mammalian cell (comprising contacting the cell with an agent that
CC   specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)
CC   a disease in a patient (comprising: (a) obtaining a biological sample
CC   from the patient; (b) contacting the sample with a probe for HM
CC   expression; (c) comparing the results with a control; and (d) determining
CC   whether the comparison indicates a likelihood disease). (M1) is useful
CC   for identifying modulators of the p53 pathway. A probe for HM expression
CC   is useful for diagnosing breast, colon, kidney, lung and ovarian cancer,
CC   in a patient, where the cancer has greater than 25 % expression level.
CC   Modulators identified by (M1) are useful in a variety of diagnostic and
CC   therapeutic applications, where disease or disorder prognosis is related
CC   to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell
CC   proliferation disorders (e.g. cancer). Another two new methods (M2 and
CC   M3) are useful for modulating the p53 pathway of a cell, thus restoring
CC   the p53 function of the cell, so that the cell undergoes normal
CC   proliferation or progression through the cell cycle. (M2) and (M3) are
CC   also useful for treating defects in the p53 pathway such as angiogenic,
CC   apoptotic or cell proliferation disorders. The present sequence
CC   represents a human p53 pathway modifying protein
XX
SQ   Sequence 316 AA;

  Query Match             65.4%;  Score 1696.5;  DB 6;  Length 316;
  Best Local Similarity   68.0%;  Pred. No. 1.8e-140;
  Matches  316;  Conservative    0;  Mismatches    0;  Indels  149;  Gaps    1;

Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQ--- 57

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Db         58 ------------------------------------------------------------ 57

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180

Db         58 ------------------------------------------------------------ 57

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240

Db         58 -------------------------EFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 91

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300

Db         92 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 151

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360

Db        152 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 211

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420

Db        212 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 271

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465

Db        272 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 316
```

```
RESULT 12
AAB67521
ID   AAB67521 standard; protein; 344 AA.
XX
AC   AAB67521;
XX
DT   29-MAY-2001  (first entry)
XX
DE   Amino acid sequence of a murine truncated parkin2 polypeptide.
XX
KW   Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW   Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW   Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
```

KW brain tumour; head trauma; stroke; vascular irregularity;
KW metabolic irregularity.
XX
OS Mus sp.
XX
PN EP1081225-A1.
XX
PD 07-MAR-2001.
XX
PF 30-AUG-1999; 99EP-00116766.
XX
PR 30-AUG-1999; 99EP-00116766.
XX
PA (BIOF-) BIOFRONTERA PHARM GMBH.
XX
PI Luebbert H;
XX
DR WPI; 2001-212797/22.
DR N-PSDB; AAF55248.
XX
PT New polynucleotides encoding mouse parkin2 protein, useful for producing
PT a transgenic non-human animal as an animal model for neurodegenerative
PT diseases.
XX
PS Claim 7; Page 40-41; 62pp; English.
XX
CC The present sequence represents a murine parkin2 polypeptide. The
CC polynucleotide sequence contains a deletion, leading to a truncated
CC protein. Mutations or deletions in the parkin2 gene cause Parkinson's
CC disease in humans. The human parkin2 gene is located in gene region
CC 6q25.2-27. Parkin2 polypeptides and polynucleotides are useful for
CC analysing neurodegenerative diseases. They are also useful for testing
CC the efficacy of the treatment of a neurodegenerative disease such as
CC Parkinson's disease, Alzheimer's disease, Huntington's disease,
CC amyotrophic lateral sclerosis, Multi-system atrophy, Wilson's disease,
CC Pick's disease, Prion disease, and secondary causes inducing Parkinson's
CC syndromes like toxins, drugs, brain tumours, head trauma, stroke,
CC vascular irregularities or metabolic irregularities, associated with a
CC less active or non-active parkin protein
XX
SQ Sequence 344 AA;

Query Match 64.7%; Score 1679.5; DB 4; Length 344;
Best Local Similarity 66.0%; Pred. No. 6.3e-139;
Matches 307; Conservative 14; Mismatches 23; Indels 121; Gaps 1;

```
Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||:|||||||||||| |||||||||||||||||||||||| |  ||
Db          1 MIVFVRFNSSYGFPVEVDSDTSILQLKEVVAKRQGVPADQLRVIFAGKELPNHLTV---- 56

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Db         57 ------------------------------------------------------------ 56

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
                                                                   |||
Db         57 ------------------------------------------------------------QGP 59

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEPFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              |||||||||||||||||| |||| ||||||||||||||:| |||:|| :| |:| ||
Db         60 SCWDDVLIPNRMSGECQSPDCPGTRAEPFFKCGAHPTSDKDTSVALNLITSNRRSIPCIA 119

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              |||||||||||||| ||||||||||||||||||||||| ||||||||||||||||||
Db        120 CTDVRSPVLVFQCNHRHVICLDCFHLYCVTRLNDRQFVHDAQLGYSLPCVAGCPNSLIKE 179

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              |||||||||||| ||||||||||||||||||||||||||||  ||||||||||||||
Db        180 LHHFRILGEEQYTRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEQGQRKVTCEGGNGLGC 239

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              || |||:|||||||:| :: | || |:||||||:||||||| ||||||||||||||
Db        240 GFVFCRDCKEAYHEGDCDSLLEPSGATSQAYRVDKRAAEQARWEEASKETIKKTTKPCPR 299

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |:||:||||||||||||||:|||||||||||||| |||||||||
Db        300 CNVPIEKNGGCMHMKCPQPQCKLEWCWNCGCEWNRACMGDHWFDV 344
```

RESULT 13
AAB67526
ID AAB67526 standard; protein; 296 AA.
XX
AC AAB67526;
XX
DT 29-MAY-2001 (first entry)
XX
DE Amino acid sequence of a murine truncated parkin2 polypeptide.
XX
KW Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW brain tumour; head trauma; stroke; vascular irregularity;
KW metabolic irregularity.
XX
OS Mus sp.
XX
PN EP1081225-A1.
XX
PD 07-MAR-2001.
XX
PF 30-AUG-1999; 99EP-00116766.
XX
PR 30-AUG-1999; 99EP-00116766.
XX
PA (BIOF-) BIOFRONTERA PHARM GMBH.
XX
PI Luebbert H;
XX
DR WPI; 2001-212797/22.
DR N-PSDB; AAF55253.
XX
PT New polynucleotides encoding mouse parkin2 protein, useful for producing
PT a transgenic non-human animal as an animal model for neurodegenerative
PT diseases.
XX
PS Claim 7; Page 44-45; 62pp; English.
XX
CC The present sequence represents a murine parkin2 polypeptide. The
CC polynucleotide sequence contains a deletion, leading to a truncated
CC protein. Mutations or deletions in the parkin2 gene cause Parkinson's
CC disease in humans. The human parkin2 gene is located in gene region
CC 6q25.2-27. Parkin2 polypeptides and polynucleotides are useful for
CC analysing neurodegenerative diseases. They are also useful for testing
CC the efficacy of the treatment of a neurodegenerative disease such as
CC Parkinson's disease, Alzheimer's disease, Huntington's disease,
CC amyotrophic lateral sclerosis, Multi-system atrophy, Wilson's disease,
CC Pick's disease, Prion disease, and secondary causes inducing Parkinson's
CC syndromes like toxins, drugs, brain tumours, head trauma, stroke,
CC vascular irregularities or metabolic irregularities, associated with a
CC less active or non-active parkin protein
XX
SQ Sequence 296 AA;

Query Match 47.6%; Score 1234.5; DB 4; Length 296;
Best Local Similarity 80.0%; Pred. No. 7.3e-100;
Matches 232; Conservative 20; Mismatches 37; Indels 1; Gaps 1;

```
Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||:|||||||||||| |||||||||||||||||||||||| |  ||||||
Db          1 MIVFVRFNSSYGFPVEVDSDTSILQLKEVVAKRQGVPADQLRVIFAGKELPNHLTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              |:||||||||||| |:  | ||:|||:|:: : |   | :|||||||||  || ||||||
Db         61 LEQQSIVHIVQRPRRRSHETNASGGDEPQSTSEGSIWESRSLTRVDLSSHTLPVDSVGLA 120
```

```
Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||| |||::|| |  |  :  ||||::|||||| :|||||||||| ||:|||||| |||
Db        121 VILDTDSKRDSEAARGPV-KPTYNSFFIYCKGPCHKVQPGKLRVQCGTCKQATLTLAQGP 179

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              |||||||||||||||||| |||| ||||||||||||||:| |||:|| :| |:| ||
Db        180 SCWDDVLIPNRMSGECQSPDCPGTRAEFFFKCGAHPTSDKDTSVALNLITSNRRSIPCIA 239

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCV 290
              |||||||||||||| |||||||||||||||||||||||| |||||||||
Db        240 CTDVRSPVLVFQCNHRHVICLDCFHLYCVTRLNDRQFVHDAQLGYSLPCV 289
```

RESULT 14
ABB65114
ID ABB65114 standard; protein; 468 AA.
XX
AC ABB65114;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 22134.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL09217.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 22134; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 468 AA;

Query Match 40.8%; Score 1058.5; DB 4; Length 468;
Best Local Similarity 42.6%; Pred. No. 4e-84;
Matches 201; Conservative 80; Mismatches 158; Indels 33; Gaps 9;

```
Qy          3 VFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCDLD 62
              ::|: |:     | ::    |  :||:|| : |:  | |::|||||||  :  |:: |||
Db         18 IYVKTNTGKTLTVNLEPQWDIKNVKELVAPQLGLQPDDLKIIFAGKELSDATTIEQCDLG 77

Qy         63 QQSIVHIVQ-RPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLAV 121
              |||::| :: ||  : |:: :   ::   :      :| : | :||
```

```
Db         78 QQSVLHAIRLRPPVQRQKIQSATLEEEEPSLSDEASKPLNETLLDLQ------------ 124

Qy        122 ILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGPS 181
               | :: | :          ::    |:|:|   | ::  |||||:|: |:    |: : |
Db        125 -LESEERLNITDEERVRAKA---HFFVHC-SQCDKLCNGKLRVRCALCKGGAFTVHRDPE 179

Qy        182 CWDDVLIPNRMSGECQSPH--CPGTS------AEFFFKCGAHPT-SDKETPVALHLIATN 232
              ||||||   |: | |:|    |   :      |||||||  | :  :|:    |:||  |
Db        180 CWDDVLKSRRIPGHCESLEVACVDNAAGDPPFAEFFFKCAEHVSGGEKDFAAPLNLIKNN 239

Qy        233 SRNITCITCTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAG 292
               :|: |: ||||   |||| | |:|| |:|||  || :|| :|||:  |  ||:||| ||
Db        240 IKNVPCLACTDVSDTVLVFPCASQHVTCIDCFRHYCRSRLGERQFMPHPDFGYTLPCPAG 299

Qy        293 CPNSLIKELHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTC 352
              | :| |:|:|||::|  |:|:|||::  || ||| ||||||:|||| ||| ||| |||||
Db        300 CEHSFIEEIHHFKLLTREEYDRYQRFATEEYVLQAGGVLCPQPGCGMGLLVEPDCRKVTC 359

Qy        353 EGGNGLGCGFAFCRECKEAYHEGEC-SAVFEASGTTTQAYRVDERAAEQARWEAASKETI 411
              :     |||: ||| | : || |||       || | :  | ||    | :|||: ||  ||
Db        360 QN----GCGYVFCRNCLQGYHIGECLPEGTGASATNSCEYTVDPNRAAEARWDEASNVTI 415

Qy        412 KKTTKPCPRCHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWF 463
              | :|||||:|  | |::|||||| | :  |  |||| |  || | ||| |||
Db        416 KVSTKPCPKCRTPTERDGGCMHMVCTRAGCGFEWCWVCQTEWTRDCMGAHWF 467
```

RESULT 15
AAB67519
ID AAB67519 standard; protein; 250 AA.
XX
AC AAB67519;
XX
DT 29-MAY-2001 (first entry)
XX
DE Amino acid sequence of a murine parkin2 polypeptide.
XX
KW Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW brain tumour; head trauma; stroke; vascular irregularity;
KW metabolic irregularity.
XX
OS Mus sp.
XX
PN EP1081225-A1.
XX
PD 07-MAR-2001.
XX
PF 30-AUG-1999; 99EP-00116766.
XX
PR 30-AUG-1999; 99EP-00116766.
XX
PA (BIOF-) BIOFRONTERA PHARM GMBH.
XX
PI Luebbert H;
XX
DR WPI; 2001-212797/22.
DR N-PSDB; AAF55246.
XX
PT New polynucleotides encoding mouse parkin2 protein, useful for producing
PT a transgenic non-human animal as an animal model for neurodegenerative
PT diseases.
XX
PS Claim 7; Page 21-22; 62pp; English.
XX
CC The present sequence represents a murine parkin2 polypeptide. Mutations
CC or deletions in the parkin2 gene cause Parkinson's disease in humans. The
CC human parkin2 gene is located in gene region 6q25.2-27. Parkin2
CC polypeptides and polynucleotides are useful for analysing
CC neurodegenerative diseases. They are also useful for testing the efficacy
CC of the treatment of a neurodegenerative disease such as Parkinson's
CC disease, Alzheimer's disease, Huntington's disease, amyotrophic lateral

```
CC   sclerosis, Multi-system atrophy, Wilson's disease, Pick's disease, Prion
CC   disease, and secondary causes inducing Parkinson's syndromes like toxins,
CC   drugs, brain tumours, head trauma, stroke, vascular irregularities or
CC   metabolic irregularities, associated with a less active or non-active
CC   parkin protein
XX
SQ   Sequence 250 AA;

  Query Match             38.6%;  Score 1002;  DB 4;  Length 250;
  Best Local Similarity   77.6%;  Pred. No. 1.7e-79;
  Matches  190;  Conservative   20;  Mismatches   35;  Indels    0;  Gaps    0;

Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||:|||||||||||| ||||||||||||||||||||||||| |  ||||||
Db          1 MIVFVRFNSSYGFPVEVDSDTSILQLKEVVAKRQGVPADQLRVIFAGKELPNHLTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              |:|||||||||||  |:  | ||:|||:|:: : |    | :|||||||||  || ||||||
Db         61 LEQQSIVHIVQRPRRRSHETNASGGDEPQSTSEGSIWESRSLTRVDLSSHTLPVDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||| |||::||  |  || :  ||||::|||||| :|||||||||| ||:|||||| |||
Db        121 VILDTDSKRDSEAARGPAVKPTYNSFFIYCKGPCHKVQPGKLRVQCGTCKQATLTLAQGP 180

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              ||||||||||||||||||| |||| ||||||||||||||:| |||:|| :| |:| ||
Db        181 SCWDDVLIPNRMSGECQSPDCPGTRAEFFFKCGAHPTSDKDTSVALNLITSNRRSIPCIA 240

Qy        241 CTDVR 245
              |||||
Db        241 CTDVR 245


Search completed: February 23, 2005, 13:50:13
Job time : 103.258 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:36:20 ; Search time 25.8801 Seconds
(without alignments)
1341.256 Million cell updates/sec

Title: US-10-622-817-9
Perfect score: 2596
Sequence: 1 MIVFVRFNSSHGFPVEVDSD.....CWNCGCEWNRVCMGDHWFDV 465

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2596	100.0	465	4	US-09-601-844B-2	Sequence 2, Appli
2	2596	100.0	465	4	US-09-949-016-6516	Sequence 6516, Ap
3	2405	92.6	437	4	US-09-601-844B-4	Sequence 4, Appli
4	2405	92.6	437	4	US-09-949-016-6853	Sequence 6853, Ap
5	1696.5	65.4	316	4	US-09-949-016-6854	Sequence 6854, Ap
6	370.5	14.3	117	4	US-09-270-767-32183	Sequence 32183, A
7	240.5	9.3	503	4	US-09-914-259-13	Sequence 13, Appl
8	222	8.6	474	4	US-09-354-221-2	Sequence 2, Appli
9	222	8.6	487	4	US-09-949-016-10111	Sequence 10111, A
10	191	7.4	506	4	US-09-248-796A-15410	Sequence 15410, A
11	167	6.4	587	1	US-08-398-008A-2	Sequence 2, Appli
12	167	6.4	587	2	US-08-893-333-2	Sequence 2, Appli
13	149	5.7	464	4	US-09-538-092-598	Sequence 598, App
14	141	5.4	328	4	US-09-248-796A-16471	Sequence 16471, A
15	133	5.1	352	3	US-08-854-764-2	Sequence 2, Appli
16	133	5.1	352	5	PCT-US95-09377-2	Sequence 2, Appli
17	131.5	5.1	1106	4	US-09-949-016-9626	Sequence 9626, Ap
18	130.5	5.0	1242	4	US-09-488-270A-2	Sequence 2, Appli
19	128.5	4.9	229	2	US-08-726-306A-23	Sequence 23, Appl
20	128.5	4.9	229	3	US-08-840-146-20	Sequence 20, Appl
21	128.5	4.9	229	3	US-09-360-220-20	Sequence 20, Appl
22	128	4.9	160	4	US-09-370-838-205	Sequence 205, App
23	128	4.9	160	4	US-09-854-133-205	Sequence 205, App
24	127	4.9	2321	4	US-09-230-652-2	Sequence 2, Appli
25	126	4.9	237	4	US-09-248-796A-19062	Sequence 19062, A
26	126	4.9	846	1	US-08-149-103-3	Sequence 3, Appli
27	126	4.9	846	1	US-08-451-883-3	Sequence 3, Appli
28	126	4.9	873	1	US-08-393-734-2	Sequence 2, Appli
29	126	4.9	873	3	US-08-894-489-2	Sequence 2, Appli
30	126	4.9	904	4	US-09-949-016-9528	Sequence 9528, Ap
31	125.5	4.8	103	2	US-08-771-201-9	Sequence 9, Appli
32	125	4.8	76	1	US-08-232-815-2	Sequence 2, Appli
33	125	4.8	76	1	US-08-350-906-2	Sequence 2, Appli
34	125	4.8	76	4	US-09-601-844B-61	Sequence 61, Appl
35	125	4.8	76	5	PCT-US95-04536-2	Sequence 2, Appli
36	125	4.8	206	4	US-09-248-796A-19430	Sequence 19430, A
37	125	4.8	1121	1	US-07-789-915A-2	Sequence 2, Appli
38	125	4.8	1121	1	US-08-005-002C-2	Sequence 2, Appli
39	125	4.8	1121	1	US-08-487-203A-2	Sequence 2, Appli
40	124.5	4.8	77	6	5510474-3	Patent No. 5510474
41	124.5	4.8	77	6	5510474-3	Patent No. 5510474
42	124.5	4.8	305	4	US-09-866-153-6	Sequence 6, Appli
43	124.5	4.8	305	4	US-09-693-467A-6	Sequence 6, Appli
44	124.5	4.8	305	4	US-09-270-976-6	Sequence 6, Appli
45	124.5	4.8	381	4	US-09-866-153-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-09-601-844B-2
; Sequence 2, Application US/09601844B
; Patent No. 6716621
; GENERAL INFORMATION:
; APPLICANT: Shimizu, No. 6716621uyoshi
; APPLICANT: Mizuno, Yoshikuni
; TITLE OF INVENTION: Isolated DNA or Gene Responsible for Parkinson's Disease
; FILE REFERENCE: 0652.2110000
; CURRENT APPLICATION NUMBER: US/09/601,844B
; CURRENT FILING DATE: 2000-08-09
; PRIOR APPLICATION NUMBER: PCT/JP99/00545
; PRIOR FILING DATE: 1999-02-09
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-601-844B-2

Query Match 100.0%; Score 2596; DB 4; Length 465;
Best Local Similarity 100.0%; Pred. No. 8.4e-251;
Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
```

```
Db        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
Db        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
```

RESULT 2
US-09-949-016-6516
; Sequence 6516, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6516
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6516

Query Match 100.0%; Score 2596; DB 4; Length 465;
Best Local Similarity 100.0%; Pred. No. 8.4e-251;
Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
Db        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
Db        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
Db        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
Db        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
Db        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
Db        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
```

RESULT 3
US-09-601-844B-4
, Sequence 4, Application US/09601844B
; Patent No. 6716621
; GENERAL INFORMATION:
; APPLICANT: Shimizu, No. 6716621uyoshi
; APPLICANT: Mizuno, Yoshikuni
; TITLE OF INVENTION: Isolated DNA or Gene Responsible for Parkinson's Disease
; FILE REFERENCE: 0652.2110000
; CURRENT APPLICATION NUMBER: US/09/601,844B
; CURRENT FILING DATE: 2000-08-09
; PRIOR APPLICATION NUMBER: PCT/JP99/00545
; PRIOR FILING DATE: 1999-02-09
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-601-844B-4

Query Match 92.6%; Score 2405; DB 4; Length 437;
Best Local Similarity 94.0%; Pred. No. 9.6e-232;
Matches 437; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

```
Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
Db        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQ-- 178
Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
Db        179 -------------------------EFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 212
Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
Db        213 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 272
Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
Db        273 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 332
Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
Db        333 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 392
Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
Db        393 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 437
```

RESULT 4
US-09-949-016-6853
; Sequence 6853, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6853
;   LENGTH: 437
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-6853

  Query Match             92.6%;  Score 2405;  DB 4;  Length 437;
  Best Local Similarity   94.0%;  Pred. No. 9.6e-232;
  Matches  437;  Conservative    0;  Mismatches    0;  Indels   28;  Gaps    1;

Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQ-- 178

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
                                          ||||||||||||||||||||||||||||||||
Db        179 -------------------------EFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 212

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 272

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 332

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 392

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |||||||||||||||||||||||||||||||||||||||||||||
Db        393 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 437


RESULT 5
US-09-949-016-6854
; Sequence 6854, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6854
;   LENGTH: 316
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-6854

  Query Match             65.4%;  Score 1696.5;  DB 4;  Length 316;
  Best Local Similarity   68.0%;  Pred. No. 4.2e-161;
  Matches  316;  Conservative    0;  Mismatches    0;  Indels  149;  Gaps    1;

Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQ--- 57

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 ------------------------------------------------------------ 57

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 ------------------------------------------------------------ 57

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
                                          ||||||||||||||||||||||||||||||||
Db         58 -------------------------EFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 91

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         92 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 151

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 211

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 271

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |||||||||||||||||||||||||||||||||||||||||||||
Db        272 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 316


RESULT 6
US-09-270-767-32183
; Sequence 32183, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32183
;   LENGTH: 117
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-270-767-32183

  Query Match             14.3%;  Score 370.5;  DB 4;  Length 117;
  Best Local Similarity   53.3%;  Pred. No. 3.8e-29;
  Matches   64;  Conservative   11;  Mismatches   40;  Indels    5;  Gaps    2;

Qy        345 PDQRKVTCEGGNGLGCGFAFCRECKEAYHEGEC-SAVFEASGTTTQAYRVDERAAEQARW 403
              || |||||:     |||: ||| | : || |||       || | :  | ||   | :|||
Db          1 PDCRKVTCKN----GCGYVFCRNCLQGYHIGECLPEGTGASATNSCEYTVDPNRAAEARW 56

Qy        404 EAASKETIKKTTKPCPRCHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWF 463
              : ||  ||| :|||||:|  | |::|||||| | :  |  |||| |  || | ||| |||
Db         57 DEASNVTIKVSTKPCPKCRTPTERDGGCMHMVCTRAGCGFEWCWVCQTEWTRDCMGAHWF 116


RESULT 7
US-09-914-259-13
; Sequence 13, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
;  APPLICANT: Makowski, Lee
;  APPLICANT:  Hyman, Paul
;  APPLICANT:  Williams, Mark
;  TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
```

```
;  FILE REFERENCE: 8471-010-999
;  CURRENT APPLICATION NUMBER: US/09/914,259
;  CURRENT FILING DATE:  2000-11-21
;  NUMBER OF SEQ ID NOS: 180
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
;   LENGTH: 503
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-914-259-13

  Query Match             9.3%;  Score 240.5;  DB 4;  Length 503;
  Best Local Similarity   27.5%;  Pred. No. 3e-15;
  Matches   72;  Conservative   42;  Mismatches  111;  Indels   37;  Gaps   13;

Qy       209 FFKCGAHPTSDKETPVALHLIATNSRNITCITCTDVRSPVLVFQCNSRHVICLDCFHLYC 268
            ||||  ||  :       |:    : |:   |  |       |  :       |  |: |:| |
Db       105 FFKC-AHVINPFNATEAIKQKTSRSQCEECEICFSQLPPDSMAGLECGHRFCMPCWHEYL 163

Qy       269 VTRLNDRQFVHDPQLGYSLPCVA-GCPNSLIKELHHFRILGEEQYN-RYQQYGAEECVLQ 326
             |:      :    || :: | | || : |: ::    :: : :   :|||       |  :
Db       164 STK------IVAEGLGQTISCAAHGC-DILVDDVTVANLVTDARVRVKYQQLITNSFV-E 215

Qy       327 MGGVL--CPRPGCG-AGLLPEPDQRKVTCEGGNGLGCGFAFCRECKEAYHEG-ECSAVFE 382
               :|  ||    |  |  :|  : |:| |:       ||  ||  | | :|:  :|  :
Db       216 CNQLLRWCPSVDCTYAVKVPYAEPRRVHCK------CGHVFCFACGENWHDPVKCRWL-- 267

Qy       383 ASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPRCHVPVEKNGGCMHMKCPQPQCR 442
             :  :  :  :| : | ||:       || |||| | :||:||| || |    |:
Db       268 -----KKWIKKCDDDSETSNWIAAN-------TKECPRCSVTIEKDGGCNHMVCKNQNCK 315

Qy       443 LEWCWNCGCEWNRVCMGDHWFD 464
              |:|| |   |     |  |::
Db       316 NEFCWVCLGSWEP--HGSSWYN 335


RESULT 8
US-09-354-221-2
; Sequence 2, Application US/09354221
; Patent No. 6699714
; GENERAL INFORMATION:
;  APPLICANT: Chang, Chawnshang
;  TITLE OF INVENTION: Androgen Receptor Coactivators
;  FILE REFERENCE: 920920.90011
;  CURRENT APPLICATION NUMBER: US/09/354,221
;  CURRENT FILING DATE: 1999-07-15
;  EARLIER APPLICATION NUMBER: US 60/100,243
;  EARLIER FILING DATE: 1998-09-14
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 474
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-354-221-2

  Query Match             8.6%;  Score 222;  DB 4;  Length 474;
  Best Local Similarity   27.8%;  Pred. No. 2e-13;
  Matches   71;  Conservative   27;  Mismatches   97;  Indels   60;  Gaps   13;

Qy       232 NSRNITC--ITCTDVRSPVLVF-QCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLP 288
             ||:    |    | : |  : | :|  ||| |   |    :: | |           :
Db       214 NSKLFLCSICFCEKLGSECMYFLEC--RHVYCKACLKDYFEIQIRDGQ----------VQ 261

Qy       289 CVAGCPN---------SLIKELHHFRILGEEQYNRYQQYGAEECVLQMGGVL-CPRPGCG 338
             |:  ||           :|||     :  | :|| :   :  :   |  |: |||| |
Db       262 CL-NCPEPKCPSVATPGQVKEL-----VEAELFARYDRLLLQSSLDLMADVVYCPRPCCQ 315

Qy       339 AGLLPEPDQRKVTCEGGNGLGCGFAFCRECKEAYH-EGECSAVFEASGTTTQAY------ 391
               :: ||       |  |    | |||| |:  ||    |    |
Db       316 LPVMQEPG-----CTMGICSSCNFAFCTLCRLTYHGVSPCKVTAEKLMDLRNEYLQADEA 370

Qy       392 -------RVDERAAEQARWEAASKETIKKTTKPCPRCHVPVEKNGGCMHMKCPQPQCRLE 444
                    |  :|  ::|  |  ||| ::| :| || |  |:|| ||  | |  |
Db       371 NKRLLDQRYGKRVIQKALEEMESKEWLEKNSKSCPCCGTPIEKLDGCNKMTC--TGCMQY 428

Qy       445 WCWNCGCEWNRVCMG 459
             :||       :|||
Db       429 FCW--------ICMG 435


RESULT 9
US-09-949-016-10111
; Sequence 10111, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10111
;   LENGTH: 487
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-10111

  Query Match             8.6%;  Score 222;  DB 4;  Length 487;
  Best Local Similarity   27.8%;  Pred. No. 2e-13;
  Matches   71;  Conservative   27;  Mismatches   97;  Indels   60;  Gaps   13;

Qy       232 NSRNITC--ITCTDVRSPVLVF-QCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLP 288
             ||:    |    | : |  : | :|  ||| |   |    :: | |           :
Db       227 NSKLFLCSICFCEKLGSECMYFLEC--RHVYCKACLKDYFEIQIRDGQ----------VQ 274

Qy       289 CVAGCPN---------SLIKELHHFRILGEEQYNRYQQYGAEECVLQMGGVL-CPRPGCG 338
             |:  ||           :|||     :  | :|| :   :  :   |  |: |||| |
Db       275 CL-NCPEPKCPSVATPGQVKEL-----VEAELFARYDRLLLQSSLDLMADVVYCPRPCCQ 328

Qy       339 AGLLPEPDQRKVTCEGGNGLGCGFAFCRECKEAYH-EGECSAVFEASGTTTQAY------ 391
               :: ||       |  |    | |||| |:  ||    |    |
Db       329 LPVMQEPG-----CTMGICSSCNFAFCTLCRLTYHGVSPCKVTAEKLMDLRNEYLQADEA 383

Qy       392 -------RVDERAAEQARWEAASKETIKKTTKPCPRCHVPVEKNGGCMHMKCPQPQCRLE 444
                    |  :|  ::|  |  ||| ::| :| || |  |:|| ||  | |  |
Db       384 NKRLLDQRYGKRVIQKALEEMESKEWLEKNSKSCPCCGTPIEKLDGCNKMTC--TGCMQY 441

Qy       445 WCWNCGCEWNRVCMG 459
             :||       :|||
Db       442 FCW--------ICMG 448


RESULT 10
US-09-248-796A-15410
; Sequence 15410, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:57:26 ; Search time 73.0113 Seconds
(without alignments)
2084.158 Million cell updates/sec

Title: US-10-622-817-9
Perfect score: 2596
Sequence: 1 MIVFVRFNSSHGFPVEVDSD.....CWNCGCEWNRVCMGDHWFDV 465

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1380268 seqs, 327241040 residues

Total number of hits satisfying chosen parameters: 1380268

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2596	100.0	465	16	US-10-473-226-2	Sequence 2, Appli
2	2596	100.0	465	16	US-10-776-604-2	Sequence 2, Appli
3	2596	100.0	465	17	US-10-839-688-9	Sequence 9, Appli
4	2584	99.5	465	16	US-10-473-226-4	Sequence 4, Appli
5	2405	92.6	437	16	US-10-776-604-4	Sequence 4, Appli
6	870	33.5	156	9	US-09-785-548-4	Sequence 4, Appli
7	544.5	21.0	386	14	US-10-239-249-2	Sequence 2, Appli
8	373	14.4	64	15	US-10-313-203-10	Sequence 10, Appl
9	323	12.4	56	15	US-10-313-203-17	Sequence 17, Appl
10	263	10.1	46	9	US-09-864-761-36750	Sequence 36750, A
11	246	9.5	525	16	US-10-437-963-199787	Sequence 199787,
12	242.5	9.3	604	15	US-10-425-114-55407	Sequence 55407, A
13	240.5	9.3	503	14	US-10-080-608A-13	Sequence 13, Appl
14	240.5	9.3	503	15	US-10-370-685-102	Sequence 102, App
15	233.5	9.0	585	15	US-10-424-599-147673	Sequence 147673,
16	224.5	8.6	648	15	US-10-425-114-63379	Sequence 63379, A
17	222	8.6	474	9	US-09-808-387-42	Sequence 42, Appl
18	222	8.6	474	14	US-10-442-754-2	Sequence 2, Appli
19	222	8.6	474	16	US-10-408-765A-572	Sequence 572, App
20	222	8.6	477	15	US-10-264-049-3019	Sequence 3019, Ap
21	219.5	8.5	1753	14	US-10-146-473-44	Sequence 44, Appl
22	219.5	8.5	1753	15	US-10-276-774-1824	Sequence 1824, Ap
23	219.5	8.5	2517	15	US-10-313-203-2	Sequence 2, Appli
24	215	8.3	32	15	US-10-313-203-18	Sequence 18, Appl
25	207.5	8.0	623	16	US-10-437-963-148905	Sequence 148905,
26	206.5	8.0	452	15	US-10-425-114-38443	Sequence 38443, A
27	205.5	7.9	449	15	US-10-425-114-39657	Sequence 39657, A
28	204	7.9	624	15	US-10-374-780A-1727	Sequence 1727, Ap
29	204	7.9	624	16	US-10-437-963-140793	Sequence 140793,
30	202.5	7.8	289	16	US-10-437-963-133385	Sequence 133385,
31	201	7.7	333	15	US-10-424-599-217796	Sequence 217796,
32	200	7.7	593	15	US-10-225-066A-854	Sequence 854, App
33	200	7.7	593	15	US-10-225-067-56	Sequence 56, Appl
34	200	7.7	593	15	US-10-374-780A-330	Sequence 330, App
35	195.5	7.5	541	15	US-10-424-599-265215	Sequence 265215,
36	193	7.4	606	16	US-10-437-963-111378	Sequence 111378,
37	192	7.4	354	15	US-10-425-114-56248	Sequence 56248, A
38	191.5	7.4	540	16	US-10-437-963-168788	Sequence 168788,
39	191	7.4	338	15	US-10-425-114-56467	Sequence 56467, A
40	184	7.1	583	16	US-10-437-963-124556	Sequence 124556,
41	182.5	7.0	522	16	US-10-437-963-184651	Sequence 184651,
42	180	6.9	250	16	US-10-437-963-152161	Sequence 152161,
43	180	6.9	394	15	US-10-425-114-54725	Sequence 54725, A
44	180	6.9	557	15	US-10-424-599-212592	Sequence 212592,
45	177.5	6.8	396	15	US-10-425-114-40231	Sequence 40231, A

ALIGNMENTS

RESULT 1
US-10-473-226-2
; Sequence 2, Application US/10473226
; Publication No. US20040198650A1
; GENERAL INFORMATION:
; APPLICANT: NsGene A/S
; TITLE OF INVENTION: Means for inhibiting proteolytical processing of Parkin
; FILE REFERENCE: 506-204-WO
; CURRENT APPLICATION NUMBER: US/10/473,226
; CURRENT FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: DK PA 2001 00525
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: US 60/281,286
; PRIOR FILING DATE: 2001-04-03
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (1)..()
; OTHER INFORMATION: Native Parkin
US-10-473-226-2

Query Match 100.0%; Score 2596; DB 16; Length 465;
Best Local Similarity 100.0%; Pred. No. 2.2e-220;
Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
Qy         121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
Db         121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
Qy         181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
Db         181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
Qy         241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
Db         241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
Qy         301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
Db         301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
Qy         361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
Db         361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
Qy         421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
Db         421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
```

RESULT 2
US-10-776-604-2
; Sequence 2, Application US/10776604
; Publication No. US20050003385A1
; GENERAL INFORMATION:
; APPLICANT: Shimizu, Nobuyoshi
; APPLICANT: Mizuno, Yoshikuni
; TITLE OF INVENTION: DNAs or Genes Participating in Parkinson's Disease
; FILE REFERENCE: 0652.2110001
; CURRENT APPLICATION NUMBER: US/10/776,604
; CURRENT FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 09/601,844
; PRIOR FILING DATE: 2000-08-09
; PRIOR APPLICATION NUMBER: PCT/JP99/00545
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: JP 10/27531
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-776-604-2

Query Match 100.0%; Score 2596; DB 16; Length 465;
Best Local Similarity 100.0%; Pred. No. 2.2e-220;
Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy           1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
Db           1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
Qy          61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
Db          61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
Qy         121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
Db         121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
Qy         181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
Db         181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
Qy         241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
Db         241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
Qy         301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
Db         301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
Qy         361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
Db         361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
Qy         421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
Db         421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
```

RESULT 3
US-10-839-688-9
; Sequence 9, Application US/10839688
; Publication No. US20050014173A1
; GENERAL INFORMATION:
; APPLICANT: Farrer, Matthew J.
; TITLE OF INVENTION: PARKINSON'S DISEASE MARKERS
; FILE REFERENCE: 07039-448001
; CURRENT APPLICATION NUMBER: US/10/839,688
; CURRENT FILING DATE: 2004-05-05
; PRIOR APPLICATION NUMBER: US 60/468,832
; PRIOR FILING DATE: 2003-05-08
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-839-688-9

Query Match 100.0%; Score 2596; DB 17; Length 465;
Best Local Similarity 100.0%; Pred. No. 2.2e-220;
Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy           1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
Db           1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
Qy          61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
Db          61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
Qy         121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
Db         121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
Qy         181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
Db         181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
Qy         241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
Db         241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
Qy         301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
Db         301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
Qy         361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
Db         361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
Qy         421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
Db         421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
```

```
RESULT 4
US-10-473-226-4
; Sequence 4, Application US/10473226
; Publication No. US20040198650A1
; GENERAL INFORMATION:
;  APPLICANT: NsGene A/S
;  TITLE OF INVENTION: Means for inhibiting proteolytical processing of Parkin
;  FILE REFERENCE: 506-204-WO
;  CURRENT APPLICATION NUMBER: US/10/473,226
;  CURRENT FILING DATE:  2003-09-29
;  PRIOR APPLICATION NUMBER: DK PA 2001 00525
;  PRIOR FILING DATE: 2001-03-29
;  PRIOR APPLICATION NUMBER: US 60/281,286
;  PRIOR FILING DATE: 2001-04-03
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 465
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: (1)..()
;   OTHER INFORMATION: Parkin with a D126E mutation
US-10-473-226-4

  Query Match             99.5%;  Score 2584;  DB 16;  Length 465;
  Best Local Similarity   99.6%;  Pred. No. 2.5e-219;
  Matches  463;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VILHTESRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETSVALHLIATNSRNITCIT 240

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||.|||||||||||||
Db        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |||||||||||||||||||||||||||||||||||||||||||||
Db        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465


RESULT 5
US-10-776-604-4
; Sequence 4, Application US/10776604
; Publication No. US20050003385A1
; GENERAL INFORMATION:
;  APPLICANT: Shimizu, Nobuyoshi
;  APPLICANT:  Mizuno, Yoshikuni
;  TITLE OF INVENTION: DNAs or Genes Participating in Parkinson's Disease
;  FILE REFERENCE: 0652.2110001
;  CURRENT APPLICATION NUMBER: US/10/776,604
;  CURRENT FILING DATE:  2004-02-12
;  PRIOR APPLICATION NUMBER: 09/601,844
;  PRIOR FILING DATE: 2000-08-09
;  PRIOR APPLICATION NUMBER: PCT/JP99/00545
;  PRIOR FILING DATE: 1999-02-09
;  PRIOR APPLICATION NUMBER: JP 10/27531
;  PRIOR FILING DATE: 1998-02-09
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 437
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-776-604-4

  Query Match             92.6%;  Score 2405;  DB 16;  Length 437;
  Best Local Similarity   94.0%;  Pred. No. 1.5e-203;
  Matches  437;  Conservative    0;  Mismatches    0;  Indels   28;  Gaps    1;

Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQ-- 178

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
                                          ||||||||||||||||||||||||||||||||
Db        179 -------------------------EFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 212

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 272

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 332

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 392

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |||||||||||||||||||||||||||||||||||||||||||||
Db        393 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 437


RESULT 6
US-09-785-548-4
; Sequence 4, Application US/09785548
; Patent No. US20020155577A1
; GENERAL INFORMATION:
;  APPLICANT: AVENTIS PHARMACEUTICALS, INC.
;  TITLE OF INVENTION: COMPOSITIONS THAT CAN BE USED FOR REGULATING THE ACTIVITY OF PARKI
;  FILE REFERENCE: ST00005
;  CURRENT APPLICATION NUMBER: US/09/785,548
;  CURRENT FILING DATE:  2001-02-20
;  NUMBER OF SEQ ID NOS: 46
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 156
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-785-548-4

  Query Match             33.5%;  Score 870;  DB 9;  Length 156;
  Best Local Similarity   99.4%;  Pred. No. 1.3e-68;
  Matches  155;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy        135 GSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGPSCWDDVLIPNRMSG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGPSCWDDVLIPNRMSG 60

Qy        195 ECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCITCTDVRSPVLVFQCN 254
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         61 ECQSPHCPGTSAEFFFKCGAHPTSDKETSVALHLIATNSRNITCITCTDVRSPVLVFQCN 120

Qy        255 SRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCV 290
              |||||||||||||||||||||||||||||||||||
Db        121 SRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCV 156
```

RESULT 7
US-10-239-249-2
; Sequence 2, Application US/10239249
; Publication No. US20030177507A1
; GENERAL INFORMATION:
; APPLICANT: HONER, MARIUS
; APPLICANT: LINK, WOLFGANG
; APPLICANT: BAUMEISTER, RALF
; TITLE OF INVENTION: NEMATODES AS MODEL ORGANISMS FOR INVESTIGATING
; TITLE OF INVENTION: NEURODEGERATIVE DISEASES AND, IN PARTICULAR, PARKINSON'S
; TITLE OF INVENTION: DISEASE, USES AND METHODS FOR FINDING SUBSTANCES AND
; TITLE OF INVENTION: GENES WHICH CAN BE USED IN TREATING SUCH DISEASES, AND
; TITLE OF INVENTION: IDENTIFICATION OF A NEMATODE GENE
; FILE REFERENCE: 02481.1804-00000
; CURRENT APPLICATION NUMBER: US/10/239,249
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: PCT/EP01/03214
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 100 14 109.9
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-239-249-2

Query Match 21.0%; Score 544.5; DB 14; Length 386;
Best Local Similarity 28.9%; Pred. No. 2.1e-39;
Matches 129; Conservative 59; Mismatches 166; Indels 93; Gaps 13;

```
Qy         22 SIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCDLDQQSIVHIVQRPWRKGQEMN 81
              :|  | : | |   :|:|:| |:| ||:|     :::  |   : : :::  :    |
Db         29 NIEDLTKDVEKLTEIPSDELEVVFCGKKLSKSTIMRDLSLTPATQIMLLRPKFNSHNENG 88

Qy         82 ATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLAVILHTDSRKDSPPAGSPAGRS 141
              ||                                  : |||             |
Db         89 AT-----------------------------------TAKITTDS------------S 99

Qy        142 IYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGPSCWDDVLIPNRMSGECQSPHC 201
              |  ||||:||  |  |: |||||  |   |    :: :   |  | |||   |:   |:
Db        100 ILGSFYVWCKN-CDDVKRGKLRVYCQKCSSTSVLVKSEPQNWSDVLKSKRIPAVCEECCT 158

Qy        202 PGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCITCTDVRSPVLVFQCNSRHVICL 261
              ||  ||| |||     :  :   ||  :  | :   |  |       :        || |: |
Db        159 PGLFAEFKFKC----LACNDPAAALTHVRGNWQMTECCVCDGKEKVIFDLGCN--HITCQ 212

Qy        262 DCFHLYCVTRLNDRQFVHDPQLGYSLPC-VAGCPNSLIKELHHFRILGEEQYNRYQQYGA 320
              ||  | :::|    ||: |  |::: |   || | :::::||| |:|:  |: ||:
Db        213 FCFRDYLLSQLERFGFVNQPPHGFTIFCPYPGC-NRVVQDVHHFHIMGQTSYSEYQRKAT 271

Qy        321 EECV-LQMGGVLCPRPGCGAGLLPEP--DQRKVTCEGGNGLGCGFAFCRECKEAYHEGEC 377
              |  : :   || ||   ||     ||  |  : |       | |:|||:|     |  |
Db        272 ERLIAVDDKGVTCPNVSCGQSFFWEPYDDDGRSQCP-----DCFFSFCRKC----FERNC 322

Qy        378 SAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPRCHVPVEKNGGCMHMKCP 437
              |   | |                                                    |
Db        323 VCQSEDDLTRT-------------------TIDATTRRCPKCHVATERNGGCAHIHC- 360
```

```
Qy        438 QPQCRLEWCWNCGCEWNRVCMGDHWFD 464
              | ::||:  |  ||   |  ||||:
Db        361 -TSCGMDWCFKCKTEWKEECQWDHWFN 386
```

RESULT 8
US-10-313-203-10
; Sequence 10, Application US/10313203
; Publication No. US20040029134A1
; GENERAL INFORMATION:
; APPLICANT: Gu, Wei
; APPLICANT: Nikolaev, Anatoly
; TITLE OF INVENTION: P53-ASSOCIATED PARKIN-LIKE CYTOPLASMIC PROTEIN, AND RELATED COMPO
; TITLE OF INVENTION: AND METHODS
; FILE REFERENCE: 68106
; CURRENT APPLICATION NUMBER: US/10/313,203
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-313-203-10

Query Match 14.4%; Score 373; DB 15; Length 64;
Best Local Similarity 100.0%; Pred. No. 3e-25;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        314 RYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGCGFAFCRECKEAYH 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGCGFAFCRECKEAYH 60

Qy        374 EGEC 377
              ||||
Db         61 EGEC 64
```

RESULT 9
US-10-313-203-17
; Sequence 17, Application US/10313203
; Publication No. US20040029134A1
; GENERAL INFORMATION:
; APPLICANT: Gu, Wei
; APPLICANT: Nikolaev, Anatoly
; TITLE OF INVENTION: P53-ASSOCIATED PARKIN-LIKE CYTOPLASMIC PROTEIN, AND RELATED COMPOS
; TITLE OF INVENTION: AND METHODS
; FILE REFERENCE: 68106
; CURRENT APPLICATION NUMBER: US/10/313,203
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-313-203-17

Query Match 12.4%; Score 323; DB 15; Length 56;
Best Local Similarity 100.0%; Pred. No. 6.7e-21;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        238 CITCTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGC 293
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CITCTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGC 56
```

RESULT 10
US-09-864-761-36750
; Sequence 36750, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:34:50 ; Search time 19.3575 Seconds
(without alignments)
2311.294 Million cell updates/sec

Title: US-10-622-817-9
Perfect score: 2596
Sequence: 1 MIVFVRFNSSHGFPVEVDSD.....CWNCGCEWNRVCMGDHWFDV 465

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	490	18.9	357	2	T23460	hypothetical prote
2	220	8.5	543	2	H84724	probable ARI-like
3	219.5	8.5	1753	2	T00350	hypothetical prote
4	205	7.9	437	2	T16477	hypothetical prote
5	205	7.9	491	2	F87793	protein C27A12.6 [
6	198.5	7.6	644	2	T02366	hypothetical prote
7	196.5	7.6	514	2	G84724	prcbable ARI-like
8	195	7.5	320	2	T45909	hypothetical prote
9	193	7.4	542	2	A84725	similar to Ariadne
10	190.5	7.3	436	2	H87793	protein C27A12.8 [
11	188	7.2	1048	2	T31653	hypothetical prote
12	186.5	7.2	324	2	T16983	hypothetical prote
13	182.5	7.0	551	2	S38086	hypothetical prote
14	180.5	7.0	594	2	T04783	hypothetical prote
15	178.5	6.9	497	2	G87793	protein C27A12.7 [
16	178.5	6.9	688	2	B86448	hypothetical prote
17	174.5	6.7	408	2	T47498	hypothetical prote
18	174	6.7	348	2	T47494	hypothetical prote
19	173	6.7	565	2	F84721	probable RING zinc
20	162.5	6.3	498	2	JC5983	protein kinase C-i
21	159	6.1	518	2	E84536	hypothetical prote
22	158.5	6.1	451	2	T16481	hypothetical prote
23	157	6.0	488	2	T29562	hypothetical prote
24	155	6.0	532	2	T04748	hypothetical prote
25	152.5	5.9	816	2	T25555	hypothetical prote
26	149	5.7	464	2	S48329	probable membrane
27	147	5.7	1209	2	T52523	hypothetical prote
28	143.5	5.5	155	2	JH0227	ubiquitin / riboso
29	142.5	5.5	155	2	JH0226	ubiquitin / riboso
30	141.5	5.5	468	2	A84601	Mutator-like trans
31	136	5.2	222	2	T47487	hypothetical prote
32	134.5	5.2	229	1	UQJNI	ubiquitin precurso
33	134	5.2	373	2	E84647	hypothetical prote
34	134	5.2	638	2	T02611	hypothetical prote
35	133	5.1	397	2	H84578	probable RING zinc
36	132.5	5.1	869	1	JC4858	VLDL receptor prec
37	132	5.1	80	2	T28305	ORF MSV144 probabl
38	132	5.1	150	2	T30390	probable ubiquitin
39	132	5.1	154	2	S55242	polyubiquitin 2 -
40	132	5.1	156	2	C86439	protein T19E23.13
41	131	5.0	534	2	S34285	polyubiquitin - to
42	130.5	5.0	305	1	S29853	polyubiquitin 4 -
43	130.5	5.0	356	1	UQUTRC	polyubiquitin / ri
44	130	5.0	155	2	S40240	ubiquitin/ribosoma
45	130	5.0	156	2	T52335	ubiquitin extensio

ALIGNMENTS

RESULT 1
T23460
hypothetical protein K08E3.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T23460
R;McMurray, A.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19743
A;Accession: T23460
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-357 <WIL>
A;Cross-references: UNIPROT:Q9XUS3; EMBL:Z81568; PIDN:CAB04599.1; GSPDB:GN00021; CESP:K08
A;Experimental source: clone K08E3
C;Genetics:
A;Gene: CESP:K08E3.7
A;Map position: 3
A;Introns: 23/3; 72/2; 218/1; 286/2; 311/2

Query Match 18.9%; Score 490; DB 2; Length 357;
Best Local Similarity 26.7%; Pred. No. 1.8e-29;
Matches 119; Conservative 53; Mismatches 154; Indels 120; Gaps 12;

```
Qy         22 SIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCDLDQQSIVHIVQRPWRKGQEMN 81
Db         29 NIEDLTKDVEKLTEIPSDELEVVFCGKKLSKSTIMRDLSLTPATQIMLLRPKFNSHNENG 88

Qy         82 ATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLAVILHTDSRKDSPPAGSPAGRS 141
Db         89 AT----------------------------------TAKITTDS------------S 99

Qy        142 IYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGPSCWDDVLIPNRMSGECQSPHC 201
Db        100 ILGSFYVWCKN-CDDVKRGKLRVYCQKCSSTSVLVKSEPQNWSDVLKSKRIPAVCEECCT 158

Qy        202 PGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCITCTDVRSPVLVFQCNSRHVICL 261
Db        159 PGLFAEFKFKC----LACNDPAAALTHVRGNWQMTECCVCDGKEKVIFDLGCN--HITCQ 212

Qy        262 DCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKELHHFRILGEEQYNRYQQYGAE 321
Db        213 FCFRV----------------------------RVVQDVHHFHIMGQTSYSEYQRKATE 243

Qy        322 ECV-LQMGGVLCPRPGCGAGLLPEP--DQRKVTCEGGNGLGCGFAFCRECKEAYHEGECS 378
Db        244 RLIAVDDKGVTCPNVSCGQSFFWEPYDDDGRSQCP-----DCFFSFCRKC----FERNCV 294

Qy        379 AVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPRCHVPVEKNGGCMHMKCPQ 438
Db        295 CQSEDDLTRT--------------------TIDATTRRCPKCHVATERNGGCAHIHC-- 331
```

```
Qy        439 PQCRLEWCWNCGCEWNRVCMGDHWFD 464
              |  ::||: |  ||   |  ||||:
Db        332 TSCGMDWCFKCKTEWKEECQWDHWFN 357
```

RESULT 2
H84724
probable ARI-like RING zinc finger protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: H84724
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L. euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84724
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-543 <STO>
A;Cross-references: UNIPROT:Q9SKC3; GB:AE002093; NID:g4887759; PIDN:AAD32295.1; GSPDB:GN
C;Genetics:
A;Gene: At2g31770
A;Map position: 2

Query Match 8.5%; Score 220; DB 2; Length 543;
Best Local Similarity 25.9%; Pred. No. 5.9e-09;
Matches 60; Conservative 33; Mismatches 83; Indels 56; Gaps 11;

```
Qy        235 NITCITCTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLND-----RQFVHDPQLGYSLPC 289
              || |  | :  :   : : :  |  |  |: |  |:: |     |     :|       |
Db        125 NIQCGICFESYTREEIARVSCGHPYCKTCWAGYITTKIEDGPGCLRVKCPEPS------C 178

Qy        290 VAGCPNSLIKELHHFRILGEEQYNRY--QQYGAEECVLQMGGVLCPRPGCGAGLL---PE 344
              |     :|:::   ::   |:|:||  : |  :   ::    || ||||  :     |
Db        179 SAAVGKDMIEDVTETKV--NEKYSRYILRSYVEDGKKIKW----CPSPGCGYAVEFGGSE 232

Qy        345 PDQRKVTCEGGNGLGCGFAFCRECKEAYHEG-ECSAVFEASGTTTQAYRVDERAAEQARW 403
              |:|       | : ||  | |  |   :|  |                    ::|
Db        233 SSSYDVSCL------CSYRFCWNCSEDAHSPVDCDTV-------------------SKW 266

Qy        404 ------EAASKETIKKTTKPCPRCHVPVEKNGGCMHMKCPQPQCRLEWCWNC 449
                    |: :|  :   :|||| |  |:||| || || |  | |  |:|| |
Db        267 IFKNQDESENKNWMLANSKPCPECKRPIEKNDGCNHMTCSAP-CGHEFCWIC 317
```

RESULT 3
T00350
hypothetical protein KIAA0708 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C;Accession: T00350
R;Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N. DNA Res. 5, 169-176, 1998
A;Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A;Reference number: Z14142; MUID:98403880; PMID:9734811
A;Accession: T00350
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1753 <ISH>
A;Cross-references: UNIPROT:Q8IWT3; EMBL:AB014608; NID:g3327229; PIDN:BAA31683.1; PID:g3
A;Experimental source: brain
C;Genetics:
A;Note: KIAA0708

Query Match 8.5%; Score 219.5; DB 2; Length 1753;
Best Local Similarity 21.4%; Pred. No. 2.2e-08;
Matches 125; Conservative 57; Mismatches 181; Indels 221; Gaps 32;

```
Qy          1 MIVFVRFNSSHGFPVEV---DSDTS-----------------------------IFQLK 27
              | : ::|| :    ||    ||| |                             : :|
Db       1012 MWLLLKFNQTEEVSVETLLKDSDLSPELLLQALVPLTSGNGPLTLHEGQDFPHGGVLRLH 1071

Qy         28 EVVAKRQG-------------VPADQLRVIFAGKELRNDWTV--------QNCDLDQQSI 66
              |   :| |                 |  |: | :    : | :  |          :   :||  :
Db       1072 EPGPQRSGEALWLIPPQAYLNVEKDEGRTLEQKRNLLSCLLVRILKAHGEKGLHIDQ--L 1129

Qy         67 VHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLAVILHTD 126
              | :|   |:||      | |       |||       : |  |: | :|    :|   :
Db       1130 VCLVLEAWQKGPNPPGTLG---HTVAGGV-----ACTSTDVLSCIL--HLLGQGYV---- 1175

Qy        127 SRKDSPP------AGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              |:|  |      |  | |     :   :|    :  :|     |          |||   | |
Db       1176 KRRDDRPQILMYAAPEPMGPCRGQADVPFCGSQSETSKPSPEAV-------ATLASLQLP 1228

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFK---CGAHPTSDKETPVALHLIATNS---- 233
                    :|   ||      |   |          | : |  || ||:| :
Db       1229 ------------AGRTMSPQ----EVEGLMKQTVRQVQETLNLEPDVAQHLLAHSHWGAE 1272

Qy        234 ----------------------------RNITCITCTDVRSPV-----LVFQCNSRHVI 259
                                         |   |  |    ||:     |  |  |
Db       1273 QLLQSYSEDPEPLLLAAGLCVHQAQAVPVRPDHCPVCV---SPLGCDDDLPSLC-CMHYC 1328

Qy        260 CLDCFHLYCVTRLNDRQFVHDPQLGYSLPC-VAGCP----NSLIKELHHFRILGEEQYNR 314
              |  |:: |  ||:       :  |  :  | :| ||      : |: :   :   |  ::
Db       1329 CKSCWNEYLTTRI-------EQNLVLNCTCPIADCPAQPTGAFIRAI----VSSPEVISK 1377

Qy        315 YQQ-----YGAEECVLQMGGVLCPRP-GCGAGLLPEPDQRKVTCEGGNGLG-----CGFA 363
              |::     |  | |              |  | ||              ::  |  | |     ||:|
Db       1378 YEKALLRGY-VESC---SNLTWCTNPQGCD----------RILCRQGLGCGTTCSKCGWA 1423

Qy        364 FCRECK--EAYHEGECSAVFEASGTTTQAYRVDERAAEQARW-----------EAASKE 409
              |  |   ||::  |                          ::|                || ||
Db       1424 SCFNCSFPEAHYPASC-------------------GHMSQWVDDGGYYDGMSVEAQSKH 1463

Qy        410 TIKKTTKPCPRCHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEW 453
                |  :| || |  |:||| ||:|| |   :|   :|| |   |
Db       1464 LAKLISKRCPSCQAPIEKNEGCLHMTC--AKCNHGFCWRCLKSW 1505
```

RESULT 4
T16477
hypothetical protein F56D2.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T16477
R;Du, Z.
submitted to the EMBL Data Library, August 1994
A;Description: The sequence of C. elegans cosmid F56D2.
A;Reference number: Z18519
A;Accession: T16477
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-437 <DUZ>
A;Cross-references: UNIPROT:Q20871; EMBL:U13644; NID:g532100; PID:g1945502; PIDN:AAB5268
A;Experimental source: strain Bristol N2; clone F56D2
C;Genetics:
A;Gene: CESP:F56D2.5
A;Map position: 3
A;Introns: 41/3; 134/3; 185/3; 223/3; 254/2; 287/2; 414/3

Query Match 7.9%; Score 205; DB 2; Length 437;
Best Local Similarity 20.2%; Pred. No. 6.3e-08;
Matches 104; Conservative 66; Mismatches 169; Indels 176; Gaps 28;

```
Qy         18 DSDTSIFQLKEVVAKRQGVPADQLRVIFAGKEL--RNDWTVQNCDLDQQSIVHIVQRPWR 75
              | |  |::|             : |  :   |:|   :||: :| ::  | |: :
Db          3 DRDLQIYEL------------EALESVLREKKLAKSSDWSDKNAEI--QGIIEV------ 42

Qy         76 KGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLAVILHTDSRKDSPPAG 135
                 | |:  :     |    |  :  |   :||       : |       | |
Db         43 --------GFDNLYDPTVTIEGTSDSGDQFHLPLDILP------PIRLKFHLPNDYPTVS 88

Qy        136 SPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGPSCWDDVLIPNRMSGE 195
              ||                  || ::     |  :|     ||         ::
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:33:04 ; Search time 88.371 Seconds
(without alignments)
2694.512 Million cell updates/sec

Title: US-10-622-817-9
Perfect score: 2596
Sequence: 1 MIVFVRFNSSHGFPVEVDSD.....CWNCGCEWNRVCMGDHWFDV 465

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2596	100.0	465	1	PRKN_HUMAN	O60260 homo sapien
2	2218	85.4	465	1	PRKN_RAT	Q9jk66 rattus norv
3	2170.5	83.6	464	1	PRKN_MOUSE	Q9wvs6 mus musculu
4	1115	43.0	203	2	Q8NI42	Q8ni42 homo sapien
5	1060	40.8	194	2	Q95M07	Q95m07 bos taurus
6	1058.5	40.8	482	2	Q95TI4	Q95ti4 drosophila
7	1058.5	40.8	482	2	Q7KTX7	Q7ktx7 drosophila
8	1049	40.4	514	2	Q7Q591	Q7q591 anopheles g
9	1024	39.4	177	2	Q8NI41	Q8ni41 homo sapien
10	957	36.9	177	2	Q8K5C2	Q8k5c2 rattus norv
11	544.5	21.0	386	2	Q9XUS3	Q9xus3 caenorhabdi
12	543.5	20.9	162	2	Q86LE7	Q86le7 drosophila
13	485	18.7	111	2	Q8VHY5	Q8vhy5 rattus norv
14	420.5	16.2	120	2	Q7TPF8	Q7tpf8 mus musculu
15	382	14.7	78	2	Q8K1N5	Q8k1n5 mus musculu
16	284	10.9	50	2	Q6Q2I6	Q6q2i6 homo sapien
17	256	9.9	520	2	Q6T486	Q6t486 dictyosteli
18	250.5	9.6	562	2	Q84RR0	Q84rr0 arabidopsis
19	249.5	9.6	492	2	Q68F65	Q68f65 xenopus lae
20	246	9.5	525	2	Q653S8	Q653s8 oryza sativ
21	240.5	9.3	503	1	ARI1_DROME	Q94981 drosophila
22	240	9.2	492	2	Q7T355	Q7t355 brachydanio
23	240	9.2	511	2	Q9VT94	Q9vt94 drosophila
24	238.5	9.2	527	2	Q7QGU7	Q7qgu7 anopheles g
25	236.5	9.1	445	2	Q8CFJ4	Q8cfj4 mus musculu
26	236.5	9.1	533	2	Q6NW85	Q6nw85 brachydanio
27	236.5	9.1	555	1	ARI1_MOUSE	Q9z1k5 mus musculu
28	236.5	9.1	557	1	ARI1_HUMAN	Q9y4x5 homo sapien
29	234.5	9.0	527	2	Q6PFJ9	Q6pfj9 brachydanio
30	232.5	9.0	492	1	ARI2_MOUSE	Q9z1k6 mus musculu
31	231	8.9	491	2	Q6GL03	Q6gl03 xenopus tro
32	231	8.9	493	1	ARI2_HUMAN	O95376 homo sapien
33	231	8.9	607	2	Q6ZIV0	Q6ziv0 oryza sativ
34	227.5	8.8	589	2	Q7XRI6	Q7xri6 oryza sativ
35	225.5	8.7	501	2	Q6CSC0	Q6csc0 kluyveromyc
36	225.5	8.7	2489	2	Q68CP2	Q68cp2 homo sapien
37	224	8.6	551	2	Q84RR1	Q84rr1 arabidopsis
38	224	8.6	552	2	Q8L829	Q8l829 arabidopsis
39	222	8.6	474	1	RN14_HUMAN	Q9ubs8 homo sapien
40	222	8.6	474	2	Q6IBV0	Q6ibv0 homo sapien
41	222	8.6	492	2	Q6RZU9	Q6rzu9 musa acumin
42	220	8.5	40	2	Q6S8G7	Q6s8g7 homo sapien
43	220	8.5	543	2	Q9SKC3	Q9skc3 arabidopsis
44	219.5	8.5	2517	1	PARC_HUMAN	Q8iwt3 homo sapien
45	219.5	8.5	2517	2	Q68D92	Q68d92 homo sapien

ALIGNMENTS

RESULT 1
PRKN_HUMAN
ID PRKN_HUMAN STANDARD; PRT; 465 AA.
AC O60260; Q8NI43; Q8NI44; Q8WW07;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Parkin (EC 6.3.2.-) (Ubiquitin E3 ligase PRKN) (Parkinson juvenile
DE disease protein 2) (Parkinson disease protein 2).
GN Name=PARK2; Synonyms=PRKN;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND INVOLVEMENT IN JUVENILE
RP PARKINSON'S DISEASE.
RC TISSUE=Fetal brain, and Skeletal muscle;
RX MEDLINE=98219084; PubMed=9560156; DOI=10.1038/33416;
RA Kitada T., Asakawa S., Hattori N., Matsumine H., Yamamura Y.,
RA Minoshima S., Yokochi M., Mizuno Y., Shimizu N.;
RT "Mutations in the parkin gene cause autosomal recessive juvenile
RT parkinsonism.";
RL Nature 392:605-608(1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).
RA D'Agata V., Scapagnini G., Cavallaro S.;
RT "Functional and molecular diversity of parkin.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 5).
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]

RP FUNCTION IN UBIQUITINATION.
RX PubMed=10973942; DOI=10.1074/jbc.C000447200;
RA Imai Y., Soda M., Takahashi R.;
RT "Parkin suppresses unfolded protein stress-induced cell death through its E3 ubiquitin-protein ligase activity.";
RL J. Biol. Chem. 275:35661-35664(2000).
RN [5]
RP FUNCTION, AND CHARACTERIZATION OF VARIANTS PD PRO-42 AND ARG-240.
RX PubMed=10888878; DOI=10.1038/77060;
RA Shimura H., Hattori N., Kubo S.-I., Mizuno Y., Asakawa S., Minoshima S., Shimizu N., Iwai K., Chiba T., Tanaka K., Suzuki T.;
RT "Familial Parkinson disease gene product, parkin, is a ubiquitin-protein ligase.";
RL Nat. Genet. 25:302-305(2000).
RN [6]
RP FUNCTION, AND INVOLVEMENT IN CANCER.
RX PubMed=12719539; DOI=10.1073/pnas.0931262100;
RA Cesari R., Martin E.S., Calin G.A., Pentimalli F., Bichi R., McAdams H., Trapasso F., Drusco A., Shimizu M., Masciullo V., D'Andrilli G., Scambia G., Picchio M.C., Alder H., Godwin A.K., Croce C.M.;
RT "Parkin, a gene implicated in autosomal recessive juvenile parkinsonism, is a candidate tumor suppressor gene on chromosome 6q25-q27.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:5956-5961(2003).
RN [7]
RP FUNCTION, INTERACTIONS WITH FBXW7 AND CUL1, AND UBIQUITINATION OF CYCLIN E.
RX PubMed=12628165;
RA Staropoli J.F., McDermott C., Martinat C., Schulman B., Demireva E., Abeliovich A.;
RT "Parkin is a component of an SCF-like ubiquitin ligase complex and protects postmitotic neurons from kainate excitotoxicity.";
RL Neuron 37:735-749(2003).
RN [8]
RP FUNCTION, UBIQUITINATION, AND S-NITROSYLATION.
RX PubMed=15105460; DOI=10.1126/science.1093891;
RA Chung K.K., Thomas B., Li X., Pletnikova O., Troncoso J.C., Marsh L., Dawson V.L., Dawson T.M.;
RT "S-nitrosylation of parkin regulates ubiquitination and compromises parkin's protective function.";
RL Science 304:1328-1331(2004).
RN [9]
RP SUBCELLULAR LOCATION.
RX PubMed=10319893;
RA Shimura H., Hattori N., Kubo S.-I., Yoshikawa M., Kitada T., Matsumine H., Asakawa S., Minoshima S., Yamamura Y., Shimizu N., Mizuno Y.;
RT "Immunohistochemical and subcellular localization of Parkin protein: absence of protein in autosomal recessive juvenile parkinsonism patients.";
RL Ann. Neurol. 45:668-672(1999).
RN [10]
RP INTERACTIONS WITH UBE2L6 AND SEPT5, AND UBIQUITINATION OF SEPT5.
RX PubMed=11078524; DOI=10.1073/pnas.240347797;
RA Zhang Y., Gao J., Chung K.K., Huang H., Dawson V.L., Dawson T.M.;
RT "Parkin functions as an E2-dependent ubiquitin-protein ligase and promotes the degradation of the synaptic vesicle-associated protein, CDCrel-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:13354-13359(2000).
RN [11]
RP UBIQUITINATION OF GPR37.
RX MEDLINE=21332586; PubMed=11439185; DOI=10.1016/S0092-8674(01)00407-X;
RA Imai Y., Soda M., Inoue H., Hattori N., Mizuno Y., Takahashi R.;
RT "An unfolded putative transmembrane polypeptide, which can lead to endoplasmic reticulum stress, is a substrate of Parkin.";
RL Cell 105:891-902(2001).
RN [12]
RP INTERACTION, UBIQUITINATION OF SNCAIP, CHARACTERIZATION OF VARIANTS AR-JP ARG-240; CYS-256; TRP-275 AND ASN-415, AND MUTAGENESIS OF CYS-337; CYS-421 AND CYS-431.
RX PubMed=11590439; DOI=10.1038/nm1001-1144;
RA Chung K.K., Zhang Y., Lim K.L., Tanaka Y., Huang H., Gao J., Ross C.A., Dawson V.L., Dawson T.M.;
RT "Parkin ubiquitinates the alpha-synuclein-interacting protein, synphilin-1: implications for Lewy-body formation in Parkinson disease.";
RL Nat. Med. 7:1144-1150(2001).
RN [13]
RP INTERACTIONS WITH STUB1 AND HSP70, AND UBIQUITINATION OF STUB1.
RX PubMed=12150907;
RA Imai Y., Soda M., Hatakeyama S., Akagi T., Hashikawa T., Nakayama K.-I., Takahashi R.;
RT "CHIP is associated with Parkin, a gene responsible for familial Parkinson's disease, and enhances its ubiquitin ligase activity.";
RL Mol. Cell 10:55-67(2002).
RN [14]
RP INTERACTION WITH SYT11, CHARACTERIZATION OF VARIANT PD GLY-289, AND MUTAGENESIS OF CYS-418.
RX PubMed=12925569; DOI=10.1093/hmg/ddg269;
RA Huynh D.P., Scoles D.R., Nguyen D., Pulst S.M.;
RT "The autosomal recessive juvenile Parkinson disease gene product, parkin, interacts with and ubiquitinates synaptotagmin XI.";
RL Hum. Mol. Genet. 12:2587-2597(2003).
RN [15]
RP UBIQUITINATION OF AN O-LINKED GLYCOSYLATED ISOFORM OF SNCAIP, SUBCELLULAR LOCATION, AND CHARACTERIZATION OF VARIANTS PD PRO-42 AND ARG-240.
RX PubMed=11431533; DOI=10.1126/science.1060627;
RA Shimura H., Schlossmacher M.G., Hattori N., Frosch M.P., Trockenbacher A., Schneider R., Mizuno Y., Kosik K.S., Selkoe D.J.;
RT "Ubiquitination of a new form of alpha-synuclein by parkin from human brain: implications for Parkinson's disease.";
RL Science 293:263-269(2001).
RN [16]
RP INVOLVEMENT IN CANCER.
RX PubMed=14614460; DOI=10.1038/sj.onc.1207072;
RA Denison S.R., Wang F., Becker N.A., Schuele B., Kock N., Phillips L.A., Klein C., Smith D.I.;
RT "Alterations in the common fragile site gene Parkin in ovarian and other cancers.";
RL Oncogene 22:8370-8378(2003).
RN [17]
RP REVIEW.
RX PubMed=15229644; DOI=10.1038/sj.embor.7400188;
RA Kahle P.J., Haass C.;
RT "How does parkin ligate ubiquitin to Parkinson's disease?";
RL EMBO Rep. 5:681-685(2004).
RN [18]
RP STRUCTURE BY NMR OF 1-76, AND INTERACTION WITH PSMD4.
RX PubMed=12634850; DOI=10.1038/sj.embor.embor764;
RA Sakata E., Yamaguchi Y., Kurimoto E., Kikuchi J., Yokoyama S., Yamada S., Kawahara H., Yokosawa H., Hattori N., Mizuno Y., Tanaka K., Kato K.;
RT "Parkin binds the Rpn10 subunit of 26S proteasomes through its ubiquitin-like domain.";
RL EMBO Rep. 4:301-306(2003).
RN [19]
RP REVIEW ON VARIANTS.
RX PubMed=14976155; DOI=10.1093/hmg/ddh089;
RA Mata I.F., Lockhart P.J., Farrer M.J.;
RT "Parkin genetics: one model for Parkinson's disease.";
RL Hum. Mol. Genet. 13:R127-R133(2004).
RN [20]
RP VARIANT AR-JP ARG-240.
RX PubMed=9731209;
RA Hattori N., Matsumine H., Asakawa S., Kitada T., Yoshino H., Elibol B., Brookes A.J., Yamamura Y., Kobayashi T., Wang M., Yoritaka A., Minoshima S., Shimizu N., Mizuno Y.;
RT "Point mutations (Thr240Arg and Gln311Stop) in the Parkin gene.";
RL Biochem. Biophys. Res. Commun. 249:754-758(1998).
RN [21]
RP ERRATUM.
RA Hattori N., Matsumine H., Asakawa S., Kitada T., Yoshino H., Elibol B., Brookes A.J., Yamamura Y., Kobayashi T., Wang M., Yoritaka A., Minoshima S., Shimizu N., Mizuno Y.;

```
RL   Biochem. Biophys. Res. Commun. 251:666-666(1998).
RN   [22]

  Query Match             100.0%;  Score 2596;  DB 1;  Length 465;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-195;
  Matches  465;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |||||||||||||||||||||||||||||||||||||||||||||
Db        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465


RESULT 2
PRKN_RAT
ID   PRKN_RAT       STANDARD;      PRT;   465 AA.
AC   Q9JK66; Q8K5C3; Q8K5C4; Q8K5C5; Q8K5C6; Q8VHY6; Q9JLL1; Q9JM64;
DT   25-OCT-2004 (Rel. 45, Created)
DT   25-OCT-2004 (Rel. 45, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Parkin (EC 6.3.2.-) (Ubiquitin E3 ligase PRKN).
GN   Name=Park2; Synonyms=Prkn;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4; 5 AND 6).
RC   STRAIN=Sprague-Dawley;
RX   MEDLINE=20153712; PubMed=10686358; DOI=10.1016/S0169-328X(99)00286-7;
RA   D'Agata V., Zhao W., Cavallaro S.;
RT   "Cloning and distribution of the rat parkin mRNA.";
RL   Brain Res. Mol. Brain Res. 75:345-349(2000).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 1), AND TISSUE SPECIFICITY.
RC   TISSUE=Hypothalamus;
RX   MEDLINE=20199898; PubMed=10737637;
RA   Gu W.-J., Abbas N., Lagunes M.Z., Parent A., Pradier L., Bohme G.A.,
RA   Agid Y., Hirsch E.C., Raisman-Vozari R., Brice A.;
RT   "Cloning of rat parkin cDNA and distribution of parkin in rat brain.";
RL   J. Neurochem. 74:1773-1776(2000).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   STRAIN=Sprague-Dawley;
RA   Hattori N., Wang M., Mizuno Y.;
RT   "The expression of parkin mRNA in developing, adult and ageing rat
RT   CNS.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   STRAIN=Sprague-Dawley; TISSUE=Brain;
RA   Soda M., Imai Y., Takahashi R.;
RT   "Molecular cloning of rat Parkin gene.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Functions within a multiprotein E3 ubiquitin ligase
CC       complex, catalyzing the covalent attachment of ubiquitin moieties
CC       onto substrate proteins. These substrates include SYT11, CCNE1,
CC       GPR37, STUB1, a 22 kDa O-linked glycosylated isoform of SNCAIP and
CC       SEPT5. May play a more general role in the ubiquitin proteasomal
CC       pathway by participating in the removal and/or detoxification of
CC       abnormally folded or damaged protein. Loss of this ubiquitin
CC       ligase activity appears to be the mechanism underlying
CC       pathogenesis of AR-JP. May protect neurons against alpha synuclein
CC       toxicity, proteasomal dysfunction, GPR37 accumulation, and
CC       kainate-induced excitotoxicity. May play a role in controlling
CC       neurotransmitter trafficking at the presynaptic terminal and in
CC       calcium-dependent exocytosis. Regulates cyclin E during neuronal
CC       apoptosis. May represent a tumor suppressor gene (By similarity).
CC   -!- PATHWAY: Ubiquitin conjugation; third step.
CC   -!- SUBUNIT: Forms an E3 ubiquitin ligase complex with UBE2L3 or
CC       UBE2L6. Part of a SCF-like complex, consisting of PARK2, CUL1 and
CC       FBXW7. Interacts with SNCAIP. Binds to the C2A and C2B domains of
CC       SYT11. Interacts and regulates the turnover of SEPT5. Part of a
CC       complex, including STUB1, HSP70, and GPR37. The amount of STUB1 in
CC       the complex increases during ER stress. STUB1 promotes the
CC       dissociation of HSP70 from PARK2 and GPR37, thus facilitating
CC       PARK2-mediated GPR37 ubiquitination. HSP70 transiently associates
CC       with unfolded GPR37 and inhibits the E3 activity of PARK2,
CC       whereas, STUB1 enhances the E3 activity of PARK2 through promotion
CC       of dissociation of HSP70 from PARK2-GPR37 complexes. Interacts
CC       with PSMD4 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. Expressed in the endoplasmic
CC       reticulum, dendrites, some presynaptic terminals and in
CC       postsynaptic densities (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=6;
CC       Name=1;
CC         IsoId=Q9JK66-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q9JK66-2; Sequence=VSP_011722, VSP_011723;
CC       Name=3;
CC         IsoId=Q9JK66-3; Sequence=VSP_011717;
CC       Name=4;
CC         IsoId=Q9JK66-4; Sequence=VSP_011718;
CC       Name=5;
CC         IsoId=Q9JK66-5; Sequence=VSP_011719;
CC       Name=6;
CC         IsoId=Q9JK66-6; Sequence=VSP_011717, VSP_011720, VSP_011721;
CC   -!- TISSUE SPECIFICITY: Largely confined to neuronal elements,
CC       including fibers and neuropil. Highly expressed at the forebrain
CC       level, in pyramidal cells of layer V, in various cortical regions
CC       and cerebellum. Expressed in the nucleus of diagonal band of
CC       Broca, nucleus basalis, bed nucleus of the stria terminalis, and
CC       olfactory tubercle. Moderate expression is seen in most neurons of
CC       the subthalamic nucleus, heart, skeletal muscle and testis.
CC       Moderate expression was found in frontal cortex, parietal cortex,
CC       cerebellum, heart, skeletal muscle and testis.
CC   -!- DOMAIN: The ubiquitin-like domain binds the PSMD4 subunit of 26S
CC       proteasomes (By similarity).
CC   -!- PTM: Auto-ubiquitinates in an E2-dependent manner leading to its
CC       own degradation (By similarity).
CC   -!- PTM: S-nitrosylated (By similarity).
CC   -!- SIMILARITY: Contains 2 IBR-type zinc fingers.
CC   -!- SIMILARITY: Contains 1 ubiquitin-like domain.
CC   -!- CAUTION: Has been said to contain 2 RING fingers, but these are
CC       not found by any domain detection methods.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF343574; AAL73348.1; -.
DR   EMBL; AF381277; AAM21452.1; -.
DR   EMBL; AF381278; AAM21453.1; -.
DR   EMBL; AF381279; AAM21454.1; -.
DR   EMBL; AF381280; AAM21455.1; -.
DR   EMBL; AF381281; AAM21456.1; -.
DR   EMBL; AF168004; AAF34874.1; -.
DR   EMBL; AF210434; AAG37013.1; -.
DR   EMBL; AF257234; AAF68666.1; -.
DR   EMBL; AB039878; BAA92431.1; -.
DR   HSSP; Q9WVS6; 1MG8.
DR   RGD; 61797; Park2.
DR   InterPro; IPR003977; parkin.
DR   InterPro; IPR000626; Ubiquitin.
DR   InterPro; IPR002867; Znf_C6HC.
DR   Pfam; PF01485; IBR; 1.
DR   Pfam; PF00240; ubiquitin; 1.
DR   PRINTS; PR01475; PARKIN.
DR   PRINTS; PR00348; UBIQUITIN.
DR   SMART; SM00647; IBR; 2.
DR   SMART; SM00213; UBQ; 1.
DR   PROSITE; PS00299; UBIQUITIN_1; FALSE_NEG.
DR   PROSITE; PS50053; UBIQUITIN_2; 1.
KW   Alternative splicing; Ligase; Metal-binding; Repeat; S-nitrosylation;
KW   Ubl conjugation; Ubl conjugation pathway; Zinc; Zinc-finger.
FT   DOMAIN        1     76       Ubiquitin-like.
FT   DOMAIN      204    238       STY11 binding 1.
FT   DOMAIN      257    293       STY11 binding 2.
FT   ZN_FING     313    377       IBR-type 1.
FT   ZN_FING     401    457       IBR-type 2.
FT   VARSPLIC      1    191       Missing (in isoform 3 and isoform 6).
FT                                /FTId=VSP_011717.
FT   VARSPLIC     57     57       Q -> QHPQDGFCHKSHLAVHNLSQQDVTQ (in
FT                                isoform 4).
FT                                /FTId=VSP_011718.
FT   VARSPLIC    179    206       Missing (in isoform 5).
FT                                /FTId=VSP_011719.
FT   VARSPLIC    390    394       AYRVD -> EDVCT (in isoform 6).
FT                                /FTId=VSP_011720.
FT   VARSPLIC    395    465       Missing (in isoform 6).
FT                                /FTId=VSP_011721.
FT   VARSPLIC    429    446       GGCMHMKCPQPQCKLEWC -> ERMHVQYTMCIPGAHGGY
FT                                (in isoform 2).
FT                                /FTId=VSP_011722.
FT   VARSPLIC    447    465       Missing (in isoform 2).
FT                                /FTId=VSP_011723.
FT   CONFLICT     24     24       F -> C (in Ref. 1; AAF34874).
FT   CONFLICT    138    138       E -> A (in Ref. 3).
FT   CONFLICT    348    348       K -> R (in Ref. 3).
SQ   SEQUENCE   465 AA;  51708 MW;  E13CF170AD6D042B CRC64;

  Query Match           85.4%;  Score 2218;  DB 1;  Length 465;
  Best Local Similarity   84.9%;  Pred. No. 6.8e-166;
  Matches  395;  Conservative   26;  Mismatches   44;  Indels    0;  Gaps    0;

Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              ||||||||||:|||||||||||||||||||||||||||||||||||||:|  ||||||
Db          1 MIVFVRFNSSYGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELQNHLTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              |:||||||||||| ||  | ||:||| |::   |   ||:|||||||| :|| ||||||
Db         61 LEQQSIVHIVQRPQRKSHETNASGGDKPQSTPEGSIWEPRSLTRVDLSSHILPADSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||| |||: ||  |  |  :  |:||:|||||| :|||||||| |||||||| |||
Db        121 VILDTDSKSDSEAARGPEAKPTYHSFFVYCKGPCHKVQPGKLRVQCGTCRQATLTLAQGP 180

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              |||||||||||||||||| |||| ||||||||||||||:| |||:||  |||:| ||
Db        181 SCWDDVLIPNRMSGECQSPDCPGTRAEFFFKCGAHPTSDKDTSVALNLITNNSRSIPCIA 240

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              |||||:|||||||| |||||||||||||||||||||||||| ||||||||||||||||||
Db        241 CTDVRNPVLVFQCNHRHVICLDCFHLYCVTRLNDRQFVHDAQLGYSLPCVAGCPNSLIKE 300

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              |||||||||||||||||||||||||||||||||||||||||||||||||  |:||||||||||
Db        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEQGQKKVTCEGGNGLGC 360

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              || |||:|||||||||| ::||||| |:|||||:|||||||| |||||||||||||||
Db        361 GFVFCRDCKEAYHEGECDSMFEASGATSQAYRVDQRAAEQARWEEASKETIKKTTKPCPR 420

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |:||:|||||||||||||:|||||||||||| |||||||||
Db        421 CNVPIEKNGGCMHMKCPQPQCKLEWCWNCGCEWNRACMGDHWFDV 465


RESULT 3
PRKN_MOUSE
ID   PRKN_MOUSE     STANDARD;      PRT;   464 AA.
AC   Q9WVS6; Q9ES22; Q9ES23;
DT   25-OCT-2004 (Rel. 45, Created)
DT   25-OCT-2004 (Rel. 45, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Parkin (EC 6.3.2.-) (Ubiquitin E3 ligase PRKN).
GN   Name=Park2; Synonyms=Prkn;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1), TISSUE SPECIFICITY, AND SUBCELLULAR
RP   LOCATION.
RC   TISSUE=Skeletal muscle;
RX   MEDLINE=20279843; PubMed=10818204;
RA   Kitada T., Asakawa S., Minoshima S., Mizuno Y., Shimizu N.;
RT   "Molecular cloning, gene expression, and identification of a splicing
RT   variant of the mouse parkin gene.";
RL   Mamm. Genome 11:417-421(2000).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS 2 AND 3), TISSUE SPECIFICITY, AND
RP   SUBCELLULAR LOCATION.
RC   STRAIN=BALB/c; TISSUE=Kidney;
RX   MEDLINE=20571394; PubMed=11122330;
RA   Stichel C.C., Augustin M., Kuehn K., Zhu X.-R., Engels P., Ullmer C.,
RA   Luebbert H.;
RT   "Parkin expression in the adult mouse brain.";
RL   Eur. J. Neurosci. 12:4181-4194(2000).
RN   [3]
RP   TISSUE SPECIFICITY AND DEVELOPMENT.
RX   PubMed=11675120;
RA   Huynh D.P., Dy M., Nguyen D., Kiehl T.-R., Pulst S.M.;
RT   "Differential expression and tissue distribution of parkin isoforms
RT   during mouse development.";
RL   Brain Res. Dev. Brain Res. 130:173-181(2001).
RN   [4]
RP   FUNCTION IN UBIQUITINATION, AND S-NITROSYLATION.
RX   PubMed=15105460; DOI=10.1126/science.1093891;
RA   Chung K.K., Thomas B., Li X., Pletnikova O., Troncoso J.C., Marsh L.,
RA   Dawson V.L., Dawson T.M.;
RT   "S-nitrosylation of parkin regulates ubiquitination and compromises
RT   parkin's protective function.";
RL   Science 304:1328-1331(2004).
RN   [5]
RP   STRUCTURE BY NMR OF 1-76.
RX   PubMed=12652124;
RA   Tashiro M., Okubo S., Shimotakahara S., Hatanaka H., Yasuda H.,
RA   Kainosho M., Yokoyama S., Shindo H.;
RT   "NMR structure of ubiquitin-like domain in PARKIN: gene product of
```

```
RT   familial Parkinson's disease.";
RL   J. Biomol. NMR 25:153-156(2003).
CC   -!- FUNCTION: Functions within a multiprotein E3 ubiquitin ligase
CC       complex, catalyzing the covalent attachment of ubiquitin moieties
CC       onto substrate proteins. These substrates include SYT11, CCNE1,
CC       GPR37, STUB1, a 22 kDa O-linked glycosylated isoform of SNCAIP and
CC       SEPT5. May play a more general role in the ubiquitin proteasomal
CC       pathway by participating in the removal and/or detoxification of
CC       abnormally folded or damaged protein. Loss of this ubiquitin
CC       ligase activity appears to be the mechanism underlying
CC       pathogenesis of AR-JP. May protect neurons against alpha synuclein
CC       toxicity, proteasomal dysfunction, GPR37 accumulation, and
CC       kainate-induced excitotoxicity. May play a role in controlling
CC       neurotransmitter trafficking at the presynaptic terminal and in
CC       calcium-dependent exocytosis. Regulates cyclin E during neuronal
CC       apoptosis. May represent a tumor suppressor gene (By similarity).
CC   -!- PATHWAY: Ubiquitin conjugation; third step.
CC   -!- SUBUNIT: Forms an E3 ubiquitin ligase complex with UPE2L3 or
CC       UBE2L6. Part of a SCF-like complex, consisting of PARK2, CUL1 and
CC       FBXW7. Interacts with SNCAIP. Binds to the C2A and C2B domains of
CC       SYT11. Interacts and regulates the turnover of SEPT5. Part of a
CC       complex, including STUB1, HSP70, and GPR37. The amount of STUB1 in
CC       the complex increases during ER stress. STUB1 promotes the
CC       dissociation of HSP70 from PARK2 and GPR37, thus facilitating
CC       PARK2-mediated GPR37 ubiquitination. HSP70 transiently associates
CC       with unfolded GPR37 and inhibits the E3 activity of PARK2,
CC       whereas, STUB1 enhances the E3 activity of PARK2 through promotion
CC       of dissociation of HSP70 from PARK2-GPR37 complexes. Interacts
CC       with PSMD4 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. Expressed in the endoplasmic
CC       reticulum, dendrites, some presynaptic terminals and in
CC       postsynaptic densities.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=1;
CC         IsoId=Q9WVS6-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q9WVS6-2; Sequence=VSP_011714, VSP_011715;
CC       Name=3;
CC         IsoId=Q9WVS6-3; Sequence=VSP_011713, VSP_011716;
CC   -!- TISSUE SPECIFICITY: Expressed in all subdivisions of the brain.
CC       Highly expressed in brainstem, cranial nerve, pontine, cerebellar
CC       nuclei, indusium griseum, nuclei reticularis, strata oriens and
CC       laccunosum moleculare of the hippocampal CA2 region. Low levels
CC       were found in the telencephalon and diencephalon. Expressed in
CC       heart, liver, skeletal muscle, kidney and testis.
CC   -!- DEVELOPMENTAL STAGE: In late 10 dpc weakly expressed in
CC       postmitotic neurons in the mantle layer of the developing nervous
CC       system. Expression increased at 11-12 dpc. At 15-16 dpc, as more
CC       specialized neurons and nonneural cells are formed, expression is
CC       more tissue specific. Expression was highest in the neurites,
CC       moderate levels were observed in the migrating postmitotic neurons
CC       in the intermediate and neopallial layers. In the diencephalon and
CC       other CNS regions, while the weakest level of expression was
CC       observed in the cell bodies. In nonneural tissues, high levels of
CC       expression were found in the muscle walls of the intestine, the
CC       blood vessels and the dermis.
CC   -!- DOMAIN: The ubiquitin-like domain binds the PSMD4 subunit of 26S
CC       proteasomes (By similarity).
CC   -!- PTM: Auto-ubiquitinates in an E2-dependent manner leading to its
CC       own degradation (By similarity).
CC   -!- PTM: S-nitrosylated.
CC   -!- SIMILARITY: Contains 2 IBR-type zinc fingers.
CC   -!- SIMILARITY: Contains 1 ubiquitin-like domain.
CC   -!- CAUTION: Has been said to contain 2 RING fingers, but these are
CC       not found by any domain detection methods.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB019558; BAA82404.1; -.
DR   EMBL; AF250293; AAG13890.1; -.
DR   EMBL; AF250294; AAG13891.1; -.
DR   EMBL; AF250295; AAG13892.1; -.
DR   PDB; 1MG8; NMR; A=1-76.
DR   MGD; MGI:1355296; Park2.
DR   InterPro; IPR003977; parkin.
DR   InterPro; IPR000626; Ubiquitin.
DR   InterPro; IPR002867; Znf_C6HC.
DR   Pfam; PF01485; IBR; 1.
DR   Pfam; PF00240; ubiquitin; 1.
DR   PRINTS; PR01475; PARKIN.
DR   PRINTS; PR00348; UBIQUITIN.
DR   SMART; SM00647; IBR; 2.
DR   SMART; SM00213; UBQ; 1.
DR   PROSITE; PS00299; UBIQUITIN_1; FALSE_NEG.
DR   PROSITE; PS50053; UBIQUITIN_2; 1.
KW   3D-structure; Alternative splicing; Ligase; Metal-binding; Repeat;
KW   S-nitrosylation; Ubl conjugation; Ubl conjugation pathway; Zinc;
KW   Zinc-finger.
FT   DOMAIN        1     76       Ubiquitin-like.
FT   DOMAIN      204    238       STY11 binding 1.
FT   DOMAIN      257    293       STY11 binding 2.
FT   ZN_FING     313    377       IBR-type 1.
FT   ZN_FING     400    456       IBR-type 2.
FT   VARSPLIC    244    261       RSPVLVFQCNHRHVICLD -> SHLPLSSGASVWTRPHLH
FT                                (in isoform 3).
FT                                /FTId=VSP_011713.
FT   VARSPLIC    245    254       SPVLVFQCNH -> FMRMSKHRTS (in isoform 2).
FT                                /FTId=VSP_011714.
FT   VARSPLIC    255    464       Missing (in isoform 2).
FT                                /FTId=VSP_011715.
FT   VARSPLIC    262    464       Missing (in isoform 3).
FT                                /FTId=VSP_011716.
FT   CONFLICT    137    137       P -> PA (in Ref. 1 and 2; AAG13890).
SQ   SEQUENCE   464 AA;  51617 MW;  5574A285A9A1B080 CRC64;

  Query Match             83.6%;  Score 2170.5;  DB 1;  Length 464;
  Best Local Similarity   83.4%;  Pred. No. 3.6e-162;
  Matches  388;  Conservative   29;  Mismatches   47;  Indels    1;  Gaps    1;

Qy          1 MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVIFAGKELRNDWTVQNCD 60
              |||||||||:|||||||||||| |||||||||||||||||||||||| |  ||||||
Db          1 MIVFVRFNSSYGFPVEVDSDTSILQLKEVVAKRQGVPADQLRVIFAGKELPNHLTVQNCD 60

Qy         61 LDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLA 120
              |:|||||||||| |:  | ||:|||:|:: : |   | :|||||||  || ||||||
Db         61 LEQQSIVHIVQRPRRRSHETNASGGDEPQSTSEGSIWESRSLTRVDLSSHTLPVDSVGLA 120

Qy        121 VILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGP 180
              ||| |||::||  |  |  :  ||||::|||||| :|||||||||| ||:|||||| |||
Db        121 VILDTDSKRDSEAARGPV-KPTYNSFFIYCKGPCHKVQPGKLRVQCGTCKQATLTLAQGP 179

Qy        181 SCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCIT 240
              |||||||||||||||||| |||| |||||||||||||:| |||:|| :| |:| ||
Db        180 SCWDDVLIPNRMSGECQSPDCPGTRAEFFFKCGAHPTSDKDTSVALNLITSNRRSIPCIA 239

Qy        241 CTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKE 300
              |||||||||||||| |||||||||||||||||||||||| ||||||||||||||||||||
Db        240 CTDVRSPVLVFQCNHRHVICLDCFHLYCVTRLNDRQFVHDAQLGYSLPCVAGCPNSLIKE 299

Qy        301 LHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGC 360
              |||||||||||| |||||||||||||||||||||||||||||||| ||||||||||||||
Db        300 LHHFRILGEEQYTRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEQGQRKVTCEGGNGLGC 359

Qy        361 GFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
              || |||:|||||||:| :: | || |:||||||:||||||||| ||||||||||||||||
Db        360 GFVFCRDCKEAYHEGDCDSLLEPSGATSQAYRVDKRAAEQARWEEASKETIKKTTKPCPR 419

Qy        421 CHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV 465
              |:||:|||||||||||||:|||||||||||| |||||||||
```

```
Db        420 CNVPIEKNGGCMHMKCPQPQCKLEWCWNCGCEWNRACMGDHWFDV 464


RESULT 4
Q8NI42
ID   Q8NI42      PRELIMINARY;      PRT;   203 AA.
AC   Q8NI42;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Parkin isoform.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   D'Agata V., Scapagnini G., Cavallaro S.;
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF381284; AAM21459.1; -.
DR   GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR   GO; GO:0006512; P:ubiquitin cycle; IEA.
DR   InterPro; IPR003977; parkin.
DR   InterPro; IPR002867; Znf_C6HC.
DR   Pfam; PF01485; IBR; 1.
DR   PRINTS; PR01475; PARKIN.
DR   SMART; SM00647; IBR; 1.
SQ   SEQUENCE   203 AA;  22192 MW;  F92E796F069DC818 CRC64;

  Query Match             43.0%;  Score 1115;  DB 2;  Length 203;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-79;
  Matches  198;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        192 MSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCITCTDVRSPVLVF 251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCITCTDVRSPVLVF 60

Qy        252 QCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKELHHFRILGEEQ 311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 QCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGCPNSLIKELHHFRILGEEQ 120

Qy        312 YNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGCGFAFCRECKEA 371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGCGFAFCRECKEA 180

Qy        372 YHEGECSAVFEASGTTTQ 389
              ||||||||||||||||||
Db        181 YHEGECSAVFEASGTTTQ 198


RESULT 5
Q95M07
ID   Q95M07      PRELIMINARY;      PRT;   194 AA.
AC   Q95M07;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Parkin (Fragment).
GN   Name=parkin gene;
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=98219084; PubMed=9560156; DOI=10.1038/33416;
RA   Kitada T., Asakawa S., Hattori N., Matsumine H., Yamamura Y.,
RA   Minoshima S., Yokochi M., Mizuno Y., Shimizu N.;
RT   "Mutations in the parkin gene cause autosomal recessive juvenile
RT   parkinsonism.";
RL   Nature 392:605-608(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=21869006; PubMed=11879804; DOI=10.1016/S0006-8993(02)02241-2;
RA   Hase A., Yamada H., Arai K., Sunada Y., Shimizu T., Matsumura K.;
RT   "Characterization of parkin in bovine peripheral nerve.";
RL   Brain Res. 930:143-149(2002).
DR   EMBL; AB060701; BAB70670.1; -.
DR   GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR   GO; GO:0006512; P:ubiquitin cycle; IEA.
DR   InterPro; IPR003977; parkin.
DR   PRINTS; PR01475; PARKIN.
FT   NON_TER       1      1
FT   NON_TER     194    194
SQ   SEQUENCE   194 AA;  21225 MW;  9B214F09FA74FAB1 CRC64;

  Query Match             40.8%;  Score 1060;  DB 2;  Length 194;
  Best Local Similarity   99.5%;  Pred. No. 2.4e-75;
  Matches  193;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         92 AGGCEREPQSLTRVDLSSSVLPGDSVGLAVILHTDSRKDSPPAGSPAGRSIYNSFYVYCK 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGGCEREPQSLTRVDLSSSVLPGDSVGLAVILHTDSRKDSPPAGSPAGRSIYNSFYVYCK 60

Qy        152 GPCQRVQPGKLRVQCSTCRQATLTLTQGPSCWDDVLIPNRMSGECQSPHCPGTSAEFFFK 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GPCQRVQPGKLRVQCSTCRQATLTLTQGPSCWDDVLIPNRMSGECQSPHCPGTSAEFFFK 120

Qy        212 CGAHPTSDKETPVALHLIATNSRNITCITCTDVRSPVLVFQCNSRHVICLDCFHLYCVTR 271
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGAHPTSDKETSVALHLIATNSRNITCITCTDVRSPVLVFQCNSRHVICLDCFHLYCVTR 180

Qy        272 LNDRQFVHDPQLGY 285
              ||||||||||||||
Db        181 LNDRQFVHDPQLGY 194


RESULT 6
Q95TI4
ID   Q95TI4      PRELIMINARY;      PRT;   482 AA.
AC   Q95TI4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   SD01679p.
GN   Name=park; Synonyms=CG10523;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA   Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA   Nunoo J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
RA   Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY058754; AAL13983.1; -.
DR   HSSP; Q862M4; 1AAR.
DR   FlyBase; FBgn0041100; park.
DR   GO; GO:0007005; P:mitochondrion organization and biogenesis; IMP.
DR   InterPro; IPR002052; N6_Mtase.
DR   InterPro; IPR003977; parkin.
DR   InterPro; IPR000626; Ubiquitin.
DR   InterPro; IPR002867; Znf_C6HC.
DR   Pfam; PF01485; IBR; 1.
DR   Pfam; PF00240; ubiquitin; 1.
DR   PRINTS; PR01475; PARKIN.
DR   PRINTS; PR00348; UBIQUITIN.
DR   SMART; SM00647; IBR; 2.
DR   SMART; SM00213; UBQ; 1.
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 25, 2005, 16:48:21 ; Search time 2103.69 Seconds
(without alignments)
10848.741 Million cell updates/sec

Title: US-10-622-817-10
Perfect score: 471
Sequence: 1 ggaagtccagcaggtagatc.....actccctgccttgtgtgtag 471

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	471	100.0	471	6	AX224425	AX224425 Sequence
2	468	99.4	2960	6	BD269536	BD269536 Mutations
3	466.4	99.0	904	9	AF381283	AF381283 Homo sapi
4	466.4	99.0	2960	6	AR492172	AR492172 Sequence
5	466.4	99.0	2960	6	AX584224	AX584224 Sequence
6	466.4	99.0	2960	6	AX584226	AX584226 Sequence
7	466.4	99.0	2960	9	AB009973	AB009973 Homo sapi
8	464.8	98.7	1575	9	BC022014	BC022014 Homo sapi
9	460.8	97.8	1157	9	AF381282	AF381282 Homo sapi
10	454	96.4	583	4	AB060701	AB060701 Bos tauru
11	353.4	75.0	1564	10	AB039878	AB039878 Rattus no
12	350.2	74.4	904	10	AF343575	AF343575 Rattus no
13	348.6	74.0	1466	10	AF168004	AF168004 Rattus no
14	348.6	74.0	1576	10	AF210434	AF210434 Rattus no
15	348.2	73.9	1157	10	AF343574	AF343574 Rattus no
16	347	73.7	1398	10	AF257234	AF257234 Rattus no
17	347	73.7	1470	10	AF381277	AF381277 Rattus no
18	347	73.7	1554	10	AF381280	AF381280 Rattus no
19	347	73.7	1578	10	AF381278	AF381278 Rattus no
20	340.2	72.2	1644	10	AB019558	AB019558 Mus muscu
21	340.2	72.2	3043	6	AX089875	AX089875 Sequence
22	340.2	72.2	3043	6	AX136840	AX136840 Sequence
23	340.2	72.2	3092	6	AX089879	AX089879 Sequence
24	340.2	72.2	3092	6	AX136844	AX136844 Sequence
25	340.2	72.2	3226	10	AF250293	AF250293 Mus muscu
26	340.2	72.2	3253	6	AX089876	AX089876 Sequence
27	340.2	72.2	3253	6	AX089878	AX089878 Sequence
28	340.2	72.2	3253	6	AX136841	AX136841 Sequence
29	340.2	72.2	3253	6	AX136843	AX136843 Sequence
30	340.2	72.2	3254	6	AX089877	AX089877 Sequence
31	340.2	72.2	3254	6	AX136842	AX136842 Sequence
32	340.2	72.2	3255	6	AX089863	AX089863 Sequence
33	340.2	72.2	3255	6	AX089880	AX089880 Sequence
34	340.2	72.2	3255	6	AX089881	AX089881 Sequence
35	340.2	72.2	3255	6	AX089882	AX089882 Sequence
36	340.2	72.2	3255	6	AX136828	AX136828 Sequence
37	340.2	72.2	3255	6	AX136845	AX136845 Sequence
38	340.2	72.2	3255	6	AX136846	AX136846 Sequence
39	340.2	72.2	3255	6	AX136847	AX136847 Sequence
40	338.6	71.9	3014	6	AX089869	AX089869 Sequence
41	338.6	71.9	3014	6	AX136834	AX136834 Sequence
42	338.4	71.8	1158	9	AF381284	AF381284 Homo sapi
43	288.4	61.2	2876	6	AR492173	AR492173 Sequence
44	265.4	56.3	3136	6	AX089872	AX089872 Sequence
45	265.4	56.3	3136	6	AX136837	AX136837 Sequence

ALIGNMENTS

RESULT 1
AX224425
LOCUS AX224425 471 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 3 from Patent WO0160857.
ACCESSION AX224425
VERSION AX224425.1 GI:15554665
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Koutnikova,H., Brice,A., Fournier,A., Pradier,L., Prades,C.,
Arnould-Reguigne,I., Rosier-Montus,M.F. and Corti,O.
TITLE Compositions useful for regulating parkin gene activity
JOURNAL Patent: WO 0160857-A 3 23-AUG-2001;
Aventis Pharma S.A. (FR) ; INSTITUT NATIONAL DE LA SANTE ET DE LA
RECHERCHE MEDICALE (INSERM) (FR)
FEATURES Location/Qualifiers
source 1. .471
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS <1. .471
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC69572.1"
/db_xref="GI:15554666"
/translation="GSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQ
GPSCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETSVALHLIATNSRNI
TCITCTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCV"
ORIGIN

Query Match 100.0%; Score 471; DB 6; Length 471;
Best Local Similarity 100.0%; Pred. No. 2.3e-143;
Matches 471; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60

Qy 61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120

```
Db         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
Db        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
Db        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
Db        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
Db        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
Db        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTGTAG 471
Db        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTGTAG 471
```

RESULT 2
BD269536

```
LOCUS       BD269536               2960 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Mutations of the parkin gene, compositions, methods and uses.
ACCESSION   BD269536
VERSION     BD269536.1  GI:33079304
KEYWORDS    JP 2002536961-A/1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2960)
  AUTHORS   Brice,A., Lucking,C., Abbas,N.E., Denefle,P., Ricard,S. and
            Bouley,S.
  TITLE     Mutations of the parkin gene, compositions, methods and uses
  JOURNAL   Patent: JP 2002536961-A 1 05-NOV-2002;
            AVENTIS PHARMA SA,INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE
            MEDICALE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002536961-A/1
            PD   05-NOV-2002
            PF   18-NOV-1999 JP 2000584062
            PR   19-NOV-1998 FR     98/14524,12-MAR-1999 US    60/124239 PR
            04-AUG-1999 FR     99/10140
            PI   ALEXIS BRICE,CHRISTOPHE LUCKING,NACER EDDINE ABBAS,PATRICE  PI
              DENEFLE,
            PI   SYLVAIN RICARD,SANDRINE BOULEY
            PC   C12N15/09,A01K67/027,C07K14/47,C07K16/18,C12N1/15,C12N1/19, PC
              C12N1/21,
            PC   C12N5/10,C12Q1/68,G01N33/15,G01N33/50,C12N15/00,C12N5/00 CC
            Mutations of the parkin gene, compositions, methods and uses FH
            Key             Location/Qualifiers
            FT   source          1. .2960
            FT                   /organism='Homo sapiens (human)'.
FEATURES            Location/Qualifiers
     source         1. .2960
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
ORIGIN
```

```
  Query Match             99.4%;  Score 468;  DB 6;  Length 2960;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-142;
  Matches  468;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
Db        504 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 563
Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
Db        564 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 623
Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
Db        624 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 683
Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
Db        684 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 743
Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
Db        744 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 803
Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
Db        804 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 863
Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
Db        864 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 923
Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
Db        924 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 971
```

RESULT 3
AF381283

```
LOCUS       AF381283               904 bp    mRNA    linear   PRI 02-MAY-2002
DEFINITION  Homo sapiens parkin isoform mRNA, complete cds, alternatively
            spliced.
ACCESSION   AF381283
VERSION     AF381283.1  GI:20385799
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 904)
  AUTHORS   D'Agata,V., Scapagnini,G. and Cavallaro,S.
  TITLE     Functional and molecular diversity of parkin
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 904)
  AUTHORS   D'Agata,V., Scapagnini,G. and Cavallaro,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAY-2001) Institute of Bioimaging and Pathophysiology
            of the Central Nervous System, Italian National Research Council
            (CNR), Piazza Roma, 2, Catania 95123, Italy
FEATURES            Location/Qualifiers
     source         1. .904
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
     CDS            74. .730
                    /note="alternatively spliced; transcript variant 7"
                    /codon_start=1
                    /product="parkin isoform"
                    /protein_id="AAM21458.1"
                    /db_xref="GI:20385800"
                    /translation="MNATGGDDPRNAAGGCEREPQSLTRVDLSSSVLPGDSVGLAVIL
                    HTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGPS
                    CWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRNITCI
                    TCTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVVCLLPGM
                    "
ORIGIN
```

```
  Query Match            99.0%;  Score 466.4;  DB 9;  Length 904;
  Best Local Similarity  99.8%;  Pred. No. 8.2e-142;
  Matches  467;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 298

Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 358

Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 418

Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 478

Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        479 CACCCCACCTCTGACAAGGAAACACCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 538

Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        539 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 598

Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        599 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 658

Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 706


RESULT 4
AR492172
LOCUS       AR492172                2960 bp    DNA     linear   PAT 15-MAY-2004
DEFINITION  Sequence 1 from patent US 6716621.
ACCESSION   AR492172
VERSION     AR492172.1  GI:47260688
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2960)
  AUTHORS   Shimizu,N. and Mizuno,Y.
  TITLE     Isolated DNA or gene responsible for Parkinson's disease
  JOURNAL   Patent: US 6716621-A 1 06-APR-2004;
FEATURES             Location/Qualifiers
     source          1. .2960
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match            99.0%;  Score 466.4;  DB 6;  Length 2960;
  Best Local Similarity  99.8%;  Pred. No. 9.6e-142;
  Matches  467;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 563

Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 623

Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 683

Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 743

Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        744 CACCCCACCTCTGACAAGGAAACACCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 803

Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 863

Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 923

Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 971


RESULT 5
AX584224
LOCUS       AX584224                2960 bp    DNA     linear   PAT 10-JAN-2003
DEFINITION  Sequence 1 from Patent WO02079459.
ACCESSION   AX584224
VERSION     AX584224.1  GI:27655742
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Jensen,P.H.
  TITLE     Means for inhibiting proteolytical processing of parkin
  JOURNAL   Patent: WO 02079459-A 1 10-OCT-2002;
            NSGENE A/S (DK)
FEATURES             Location/Qualifiers
     source          1. .2960
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     mRNA            1. .2960
ORIGIN

  Query Match            99.0%;  Score 466.4;  DB 6;  Length 2960;
  Best Local Similarity  99.8%;  Pred. No. 9.6e-142;
  Matches  467;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 563

Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 623

Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 683

Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 743

Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        744 CACCCCACCTCTGACAAGGAAACACCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 803

Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 863
```

```
Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 923

Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 971


RESULT 6
AX584226
LOCUS       AX584226               2960 bp    DNA     linear   PAT 10-JAN-2003
DEFINITION  Sequence 3 from Patent WO02079459.
ACCESSION   AX584226
VERSION     AX584226.1  GI:27655743
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Jensen,P.H.
  TITLE     Means for inhibiting proteolytical processing of parkin
  JOURNAL   Patent: WO 02079459-A 3 10-OCT-2002;
            NSGENE A/S (DK)
FEATURES             Location/Qualifiers
     source          1. .2960
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     mRNA            1. .2960
ORIGIN

  Query Match             99.0%;  Score 466.4;  DB 6;  Length 2960;
  Best Local Similarity   99.8%;  Pred. No. 9.6e-142;
  Matches  467;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 563

Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 623

Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 683

Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 743

Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        744 CACCCCACCTCTGACAAGGAAACACCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 803

Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 863

Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 923

Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 971


RESULT 7
AB009973
LOCUS       AB009973               2960 bp    mRNA    linear   PRI 14-APR-2000
DEFINITION  Homo sapiens mRNA for Parkin, complete cds.
ACCESSION   AB009973
VERSION     AB009973.1  GI:3063387
KEYWORDS    Parkin.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kitada,T., Asakawa,S., Hattori,N., Matsumine,H., Yamamura,Y.,
            Minoshima,S., Yokochi,M., Mizuno,Y. and Shimizu,N.
  TITLE     Mutations in the parkin gene cause autosomal recessive juvenile
            parkinsonism
  JOURNAL   Nature 392 (6676), 605-608 (1998)
  MEDLINE   98219084
   PUBMED   9560156
REFERENCE   2  (bases 1 to 2960)
  AUTHORS   Shimizu,N., Kitada,T. and Asakawa,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-DEC-1997) Nobuyoshi Shimizu, Keio University School
            of Medicine, Department of Molecular Biology; 35 Shinanomachi,
            Shinjuku-ku, Tokyo 160-8582, Japan
            (E-mail:shimizu@dmb.med.keio.ac.jp, Tel:81-3-3351-2370(ex.2720),
            Fax:81-3-3351-2370)
FEATURES             Location/Qualifiers
     source          1. .2960
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6q25.2-q27"
     gene            1. .2960
                     /gene="parkin"
     exon            1. .108
                     /gene="parkin"
                     /number=1
     CDS             102. .1499
                     /gene="parkin"
                     /codon_start=1
                     /product="Parkin"
                     /protein_id="BAA25751.1"
                     /db_xref="GI:3063388"
                     /translation="MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVI
                     FAGKELRNDWTVQNCDLDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSL
                     TRVDLSSSVLPGDSVGLAVILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPG
                     KLRVQCSTCRQATLTLTQGPSCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTS
                     DKETPVALHLIATNSRNITCITCTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQ
                     FVHDPQLGYSLPCVAGCPNSLIKELHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPR
                     PGCGAGLLPEPDQRKVTCEGGNGLGCGFAFCRECKEAYHEGECSAVFEASGTTTQAYR
                     VDERAAEQARWEAASKETIKKTTKPCPRCHVPVEKNGGCMHMKCPQPQCRLEWCWNCG
                     CEWNRVCMGDHWFDV"
     exon            109. .272
                     /gene="parkin"
                     /number=2
     exon            273. .513
                     /gene="parkin"
                     /number=3
     exon            514. .635
                     /gene="parkin"
                     /number=4
     exon            636. .719
                     /gene="parkin"
                     /number=5
     exon            720. .835
                     /gene="parkin"
                     /number=6
     exon            836. .972
                     /gene="parkin"
                     /number=7
     exon            973. .1034
                     /gene="parkin"
                     /number=8
     exon            1035. .1184
```

```
                     /gene="parkin"
                     /number=9
     exon            1185. .1268
                     /gene="parkin"
                     /number=10
     exon            1269. .1386
                     /gene="parkin"
                     /number=11
     exon            1387. .2960
                     /gene="parkin"
                     /number=12
ORIGIN

  Query Match           99.0%;  Score 466.4;  DB 9;  Length 2960;
  Best Local Similarity  99.8%;  Pred. No. 9.6e-142;
  Matches  467;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 563

Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 623

Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 683

Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 743

Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        744 CACCCCACCTCTGACAAGGAAACACCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 803

Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 863

Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 923

Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 971


RESULT 8
BC022014
LOCUS       BC022014                1575 bp    mRNA    linear   PRI 06-OCT-2003
DEFINITION  Homo sapiens Parkinson disease (autosomal recessive, juvenile) 2,
            parkin, mRNA (cDNA clone MGC:26491 IMAGE:4824892), complete cds.
ACCESSION   BC022014
VERSION     BC022014.2  GI:34191069
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1575)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1575)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-JAN-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     On Aug 25, 2003 this sequence version replaced gi:18314632.
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI) &  Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Institute for Systems Biology
            http://www.systemsbiology.org
            contact: amadan@systemsbiology.org
            Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha
            Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 33 Row: j Column: 6
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 4758883.
FEATURES             Location/Qualifiers
     source          1. .1575
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:26491 IMAGE:4824892"
                     /tissue_type="Testis"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Vector: pBluescript"
     gene            1. .1575
                     /gene="PARK2"
                     /note="synonyms: AR-JP, PDJ, PRKN"
                     /db_xref="LocusID:5071"
                     /db_xref="MIM:602544"
     CDS             99. .1262
                     /gene="PARK2"
                     /codon_start=1
                     /product="PARK2 protein"
                     /protein_id="AAH22014.1"
                     /db_xref="GI:18314633"
                     /db_xref="LocusID:5071"
                     /db_xref="MIM:602544"
                     /translation="MIVFVRFNSSHGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVI
                     FAGKELRNDWTVQNCDLDQQSIVHIVQRPWRKGQEMNATGGDDPRNAAGGCEREPQSL
                     TRVDLSSSVLPGDSVGLAVILHTDSRKDSPPAGSPAGRSIYNSFYVYCKGPCQRVQPG
                     KLRVQCSTCRQATLTLTQGPSCWDDVLIPNRMSGECQSPHCPGTSAEFFFKCGAHPTS
                     DKETSVALHLIATNSRNITCITCTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQ
                     FVHDPQLGYSLPCVGTGDTVVLRGALGGFRRGVAGCPNSLIKELHHFRILGEEQYNRY
                     QQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGCGYGQRRTK"
     misc_feature    99. .317
                     /gene="PARK2"
                     /note="ubiquitin; Region: Ubiquitin family. This family
                     contains a number of ubiquitin-like proteins: SUMO (smt3
```

homologue), Nedd8, Elongin B, Rub1"
/db_xref="CDD:pfam00240"
ORIGIN

Query Match 98.7%; Score 464.8; DB 9; Length 1575;
Best Local Similarity 99.6%; Pred. No. 3e-141;
Matches 466; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        501 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGTAAAGGCCCCTGT 560

Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        561 CAAAGAGTGCAGCCGGGGAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 620

Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        621 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 680

Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        681 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 740

Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        741 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 800

Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        801 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 860

Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        861 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 920

Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        921 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 968
```

RESULT 9
AF381282
LOCUS AF381282 1157 bp mRNA linear PRI 02-MAY-2002
DEFINITION Homo sapiens parkin isoform mRNA, complete cds, alternatively spliced.
ACCESSION AF381282
VERSION AF381282.1 GI:20385797
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1157)
AUTHORS D'Agata,V., Scapagnini,G. and Cavallaro,S.
TITLE Functional and molecular diversity of parkin
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1157)
AUTHORS D'Agata,V., Scapagnini,G. and Cavallaro,S.
TITLE Direct Submission
JOURNAL Submitted (01-MAY-2001) Institute of Bioimaging and Pathophysiology of the Central Nervous System, Italian National Research Council (CNR), Piazza Roma, 2, Catania 95123, Italy
FEATURES Location/Qualifiers
source 1. .1157
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
CDS 333. .1157
/note="alternatively spliced; transcript variant 6"
/codon_start=1
/product="parkin isoform"
/protein_id="AAM21457.1"
/db_xref="GI:20385798"
/translation="MSGECQSPHCPGTSAEFFFKCGAHPTSDKETPVALHLIATNSRN
ITCITCTDVRSPVLVFQCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGYSLPCVAGC
PNSLIKELHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEPDQRKVT
CEGGNGLGCGFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDERAAEQARWEAASKE
TIKKTTKPCPRCHVPVEKNGGCMHMKCPQPQCRLEWCWNCGCEWNRVCMGDHWFDV"
ORIGIN

Query Match 97.8%; Score 460.8; DB 9; Length 1157;
Best Local Similarity 99.6%; Pred. No. 5.9e-140;
Matches 462; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          5 GTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGTCAAA 64
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 GTGCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGTCAAA 225

Qy         65 GAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTCACCT 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 GAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTCACCT 285

Qy        125 TGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGTGAAT 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 TGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGTGAAT 345

Qy        185 GCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCACACC 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 GCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCACACC 405

Qy        245 CCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGGAACA 304
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        406 CCACCTCTGACAAGGAAACACCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGGAACA 465

Qy        305 TCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAACTCCC 364
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 TCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAACTCCC 525

Qy        365 GCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGATCGGC 424
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 GCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGATCGGC 585

Qy        425 AGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              ||||||||||||||||||||||||||||||||||||||||||||
Db        586 AGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 629
```

RESULT 10
AB060701
LOCUS AB060701 583 bp mRNA linear MAM 20-APR-2002
DEFINITION Bos taurus mRNA for parkin, partial cds.
ACCESSION AB060701
VERSION AB060701.1 GI:16444927
KEYWORDS .
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
REFERENCE 1
AUTHORS Kitada,T., Asakawa,S., Hattori,N., Matsumine,H., Yamamura,Y., Minoshima,S., Yokochi,M., Mizuno,Y. and Shimizu,N.
TITLE Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism
JOURNAL Nature 392 (6676), 605-608 (1998)
MEDLINE 98219084
PUBMED 9560156
REFERENCE 2
AUTHORS Hase,A., Yamada,H., Arai,K., Sunada,Y., Shimizu,T. and Matsumura,K.
TITLE Characterization of parkin in bovine peripheral nerve
JOURNAL Brain Res. 930 (1-2), 143-149 (2002)
MEDLINE 21869006
PUBMED 11879804
REFERENCE 3 (bases 1 to 583)
AUTHORS Hase,A., Yamada,H., Arai,K., Shimizu,T. and Matsumura,K.

```
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-2001) Asako Hase, Teikyo University School of
            Medicine, Dept. of Neurology and Neuroscience; 2-11-1, Kaga,
            Itabashi-ku, Tokyo 173-8605, Japan (E-mail:hase@med.teikyo-u.ac.jp,
            Tel:81-3-3964-1211(ex.1916), Fax:81-3-3964-6394)
FEATURES             Location/Qualifiers
     source          1. .583
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9913"
                     /tissue_type="brain"
                     /note="Identical sequence is also seen in peripheral
                     nerve"
     gene            1. .583
                     /gene="parkin gene"
     CDS             <1. .>583
                     /gene="parkin gene"
                     /codon_start=1
                     /evidence=experimental
                     /product="parkin"
                     /protein_id="BAB70670.1"
                     /db_xref="GI:16444928"
                     /translation="AGGCEREPQSLTRVDLSSSVLPGDSVGLAVILHTDSRKDSPPAG
                     SPAGRSIYNSFYVYCKGPCQRVQPGKLRVQCSTCRQATLTLTQGPSCWDDVLIPNRMS
                     GECQSPHCPGTSAEFFFKCGAHPTSDKETSVALHLIATNSRNITCITCTDVRSPVLVF
                     QCNSRHVICLDCFHLYCVTRLNDRQFVHDPQLGY"
ORIGIN

  Query Match             96.4%;  Score 454;  DB 4;  Length 583;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-138;
  Matches  454;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 189

Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 249

Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 309

Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 369

Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 429

Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 489

Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 549

Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACT 454
              ||||||||||||||||||||||||||||||||||
Db        550 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACT 583


RESULT 11
AB039878
LOCUS       AB039878               1564 bp    mRNA    linear   ROD 16-APR-2004
DEFINITION  Rattus norvegicus mRNA for Parkin, complete cds.
ACCESSION   AB039878
VERSION     AB039878.1  GI:7229096
KEYWORDS    Parkin.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1
  AUTHORS   Hattori,N., Shimura,H., Kubo,S., Kitada,T., Wang,M., Asakawa,S.,
            Minashima,S., Shimizu,N., Suzuki,T., Tanaka,K. and Mizuno,Y.
  TITLE     Autosomal recessive juvenile parkinsonism: A key to understanding
            nigral degeneration in sporadic Parkinson's disease
  JOURNAL   Neuropathology 20 s1, 85-90 (2000)
REFERENCE   2  (bases 1 to 1564)
  AUTHORS   Hattori,N., Wang,M. and Mizuno,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-MAR-2000) Nobutaka Hattori, Juntendo University,
            Dept. of Neurology; Hongo 2-1-1, Bunkyo, Tokyo 113-8421, Japan
            (E-mail:nhattori@med.juntendo.ac.jp, Tel:81-3-3813-3111(ex.3328),
            Fax:81-3-3813-7440)
FEATURES             Location/Qualifiers
     source          1. .1564
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Sragque-Dawley"
                     /db_xref="taxon:10116"
                     /sex="male"
                     /clone_lib="recombinant bacterial type"
                     /dev_stage="adult"
                     /tissue_lib="Brain"
                     /note="vector:pAP3neo"
     gene            1. .1564
                     /gene="parkin"
     CDS             71. .1468
                     /gene="parkin"
                     /codon_start=1
                     /product="Parkin"
                     /protein_id="BAA92431.1"
                     /db_xref="GI:7229097"
                     /translation="MIVFVRFNSSYGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVI
                     FAGKELQNHLTVQNCDLEQQSIVHIVQRPQRKSHETNASGGDKPQSTPEGSIWEPRSL
                     TRVDLSSHILPADSVGLAVILDTDSKSDSEAARGPAAKPTYHSFFVYCKGPCHKVQPG
                     KLRVQCGTCRQATLTLAQGPSCWDDVLIPNRMSGECQSPDCPGTRAEFFFKCGAHPTS
                     DKDTSVALNLITNNSRSIPCIACTDVRNPVLVFQCNHRHVICLDCFHLYCVTRLNDRQ
                     FVHDAQLGYSLPCVAGCPNSLIKELHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPR
                     PGCGAGLLPEQGQRKVTCEGGNGLGCGFVFCRDCKEAYHEGECDSMFEASGATSQAYR
                     VDQRAAEQARWEEASKETIKKTTKPCPRCNVPIEKNGGCMHMKCPQPQCKLEWCWNCG
                     CEWNRACMGDHWFDV"
ORIGIN

  Query Match             75.0%;  Score 353.4;  DB 10;  Length 1564;
  Best Local Similarity   84.8%;  Pred. No. 1.4e-104;
  Matches  396;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy          2 GAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGTC 61
              || |||||||| || | | | |||| ||||||||| ||| || ||||||||||||||| |
Db        474 GAGGTCCAGCAGCTAAACCCACCTACCACAGCTTTTTTGTCTACTGCAAAGGCCCCTGCC 533

Qy         62 AAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTCA 121
              | |  || ||||| || |||||| | || |||||| |||||||||| || ||||| ||||
Db        534 ACAAGGTCCAGCCTGGGAAACTCCGAGTTCAGTGCGGCACCTGCAGACAAGCAACCCTCA 593

Qy        122 CCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGTG 181
              ||||| ||||||| |||||||||||||||||||| |||||||||||||||||||||||| |
Db        594 CCTTGGCCCAGGGCCCATCTTGCTGGGATGATGTCTTAATTCCAAACCGGATGAGTGGAG 653

Qy        182 AATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCAC 241
              | |||||||| ||| |||||||||||||| || || |||||||||||||||||||||||||
Db        654 AGTGCCAATCTCCAGACTGCCCTGGGACAAGAGCTGAATTTTTCTTTAAATGTGGAGCAC 713

Qy        242 ACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGGA 301
              |||| ||||| |||||||| |||||||||||||| |||||||| | |  || || || |
Db        714 ACCCAACCTCAGACAAGGACACATCAGTAGCTTTGAACCTGATCACCAACAACAGCCGCA 773

Qy        302 ACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAACT 361
              |||| | |||||  ||||||| || ||||||| ||| ||| |||| ||||| || |||
```

```
Db        774 GCATCCCCTGCATCGCGTGCACGGATGTCAGGAACCCTGTCTTGGTCTTCCAATGTAACC 833

Qy        362 CCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGATC 421
              |||||||||||| || || ||||| |||||||| |||||||| ||||| ||||| ||||
Db        834 ACCGCCACGTGATCTGTTTGGACTGCTTCCACTTGTACTGTGTCACAAGGCTCAACGATC 893

Qy        422 GGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              |||||||||| |||||| ||||||||||||||||||||||| ||||||
Db        894 GGCAGTTTGTCCACGACGCTCAACTTGGCTACTCCCTGCCGTGTGTG 940


RESULT 12
AF343575
LOCUS       AF343575                904 bp    mRNA    linear   ROD 02-FEB-2002
DEFINITION  Rattus norvegicus parkin transcript variant 7 mRNA, complete cds.
ACCESSION   AF343575
VERSION     AF343575.1  GI:18478869
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 904)
  AUTHORS   D'Agata,V.M., Scapagnini,G. and Cavallaro,S.
  TITLE     Functional and molecular diversity of parkin in the rat brain
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 904)
  AUTHORS   D'Agata,V.M., Scapagnini,G. and Cavallaro,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JAN-2001) Institute of Bioimaging and Pathophysiology
            of the Central Nervous System, Italian National Research Council
            (CNR), Piazza Roma, 2, Catania 95123, Italy
FEATURES             Location/Qualifiers
     source          1. .904
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
     CDS             410. .745
                     /codon_start=1
                     /product="parkin transcript variant 7"
                     /protein_id="AAL73349.1"
                     /db_xref="GI:18478870"
                     /translation="MSGECQSPDCPGTRAEFFFKCGAHPTSDKDTSVALNLITNNSRS
                     IPCIACTDVRNPVLVFQCNHRHVICLDCFHLYCVTRLNDRQFVHDAQLGYSLPCVVRF
                     LPGLQGSIP"
ORIGIN

  Query Match           74.4%;  Score 350.2;  DB 10;  Length 904;
  Best Local Similarity   84.4%;  Pred. No. 1.5e-103;
  Matches  394;  Conservative    0;  Mismatches   73;  Indels    0;  Gaps    0;

Qy          2 GAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGTC 61
              || |||||||||| || | | | |||| |||||||||| ||| || |||||||||||||| |
Db        240 GAGGTCCAGCAGCTAAACCCACCTACCACAGCTTTTTTGTCTACTGCAAAGGCCCCTGCC 299

Qy         62 AAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTCA 121
              | |  || ||||| || |||||| | || |||||| |||||||||| || |||||| ||||
Db        300 ACAAGGTCCAGCCTGGGAAACTCCGAGTTCAGTGCGGCACCTGCAGACAAGCAACCCTCA 359

Qy        122 CCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGTG 181
              ||||| |||||||| |||||||||||||||||||| |||||||||||||||||||||| |
Db        360 CCTTGGCCCAGGGCCCATCTTGCTGGGATGATGTCTTAATTCCAAACCGGATGAGTGGAG 419

Qy        182 AATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCAC 241
              | |||||||| ||| |||||||||||||| || || ||||||||||||||||||||||||
Db        420 AGTGCCAATCTCCAGACTGCCCTGGGACAAGAGCTGAATTTTTCTTTAAATGTGGAGCAC 479

Qy        242 ACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGGA 301
              |||| ||||| |||||||| ||||||||||||||| |||||||| | |  || || || |
Db        480 ACCCAACCTCAGACAAGGACACATCAGTAGCTTTGAACCTGATCACCAACAACAGCCGCA 539

Qy        302 ACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAACT 361
              |||| | |||||  ||||||| || ||||||| ||| ||| |||| ||||| || |||
Db        540 GCATCCCCTGCATCGCGTGCACGGATGTCAGGAACCCTGTCTTGGTCTTCCAATGTAACC 599

Qy        362 CCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGATC 421
              |||||||||||| || || ||||| |||||||| |||||||| ||||| ||||| ||||
Db        600 ACCGCCACGTGATCTGTTTGGACTGCTTCCACTTGTACTGTGTCACAAGGCTCAACGATC 659

Qy        422 GGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              |||||||||| |||||| |||| |||||||||||| ||||| ||||||
Db        660 GGCAGTTTGTCCACGACGCTCAGCTTGGCTACTCGCTGCCGTGTGTG 706


RESULT 13
AF168004
LOCUS       AF168004                1466 bp    mRNA    linear   ROD 19-FEB-2000
DEFINITION  Rattus norvegicus Parkin mRNA, partial cds.
ACCESSION   AF168004
VERSION     AF168004.1  GI:7001383
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1466)
  AUTHORS   D'Agata,V., Zhao,W. and Cavallaro,S.
  TITLE     Cloning and distribution of the rat parkin mRNA
  JOURNAL   Brain Res. Mol. Brain Res. 75 (2), 345-349 (2000)
  MEDLINE   20153712
   PUBMED   10686358
REFERENCE   2  (bases 1 to 1466)
  AUTHORS   D'Agata,V. and Cavallaro,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUL-1999) Institute of Bioimaging and Pathophysiology
            of the Central Nervous System, Italian National Research Council,
            Piazza Roma, 2, Catania 95123, Italy
FEATURES             Location/Qualifiers
     source          1. .1466
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /tissue_type="brain"
     CDS             91. .>1466
                     /codon_start=1
                     /product="Parkin"
                     /protein_id="AAF34874.1"
                     /db_xref="GI:7001384"
                     /translation="MIVFVRFNSSYGFPVEVDSDTSICQLKEVVAKRQGVPADQLRVI
                     FAGKELQNHLTVQNCDLEQQSIVHIVQRPQRKSHETNASGGDKPQSTPEGSIWEPRSL
                     TRVDLSSHILPADSVGLAVILDTDSKSDSEAARGPEAKPTYHSFFVYCKGPCHKVQPG
                     KLRVQCGTCRQATLTLAQGPSCWDDVLIPNRMSGECQSPDCPGTRAEFFFKCGAHPTS
                     DKDTSVALNLITNNSRSIPCIACTDVRNPVLVFQCNHRHVICLDCFHLYCVTRLNDRQ
                     FVHDAQLGYSLPCVAGCPNSLIKELHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPR
                     PGCGAGLLPEQGQKKVTCEGGNGLGCGFVFCRDCKEAYHEGECDSMFEASGATSQAYR
                     VDQRAAEQARWEEASKETIKKTTKPCPRCNVPIEKNGGCMHMKCPQPQCKLEWCWNCG
                     CEWNRACMG"
ORIGIN

  Query Match           74.0%;  Score 348.6;  DB 10;  Length 1466;
  Best Local Similarity   84.2%;  Pred. No. 5.3e-103;
  Matches  393;  Conservative    0;  Mismatches   74;  Indels    0;  Gaps    0;

Qy          2 GAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGTC 61
              || |||||| || || | | | |||| |||||||||| ||| || |||||||||||||| |
Db        494 GAGGTCCAGAAGCTAAACCCACCTACCACAGCTTTTTTGTCTACTGCAAAGGCCCCTGCC 553

Qy         62 AAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTCA 121
              | |  || ||||| || |||||| | || |||||| |||||||||| || |||||| ||||
Db        554 ACAAGGTCCAGCCTGGGAAACTCCGAGTTCAGTGCGGCACCTGCAGACAAGCAACCCTCA 613

Qy        122 CCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGTG 181
              ||||| |||||||| |||||||||||||||||||| |||||||||||||||||||||| |
Db        614 CCTTGGCCCAGGGCCCATCTTGCTGGGATGATGTCTTAATTCCAAACCGGATGAGTGGAG 673
```

```
Qy        182 AATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCAC 241
              | ||||||||| ||| |||||||||||||| || || ||||||||||||||||||||||||
Db        674 AGTGCCAATCTCCAGACTGCCCTGGGACAAGAGCTGAATTTTTCTTTAAATGTGGAGCAC 733

Qy        242 ACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGGA 301
              |||| ||||| |||||||| |||||||||||||| |||||||| | |  || || || |
Db        734 ACCCAACCTCAGACAAGGACACATCAGTAGCTTTGAACCTGATCACCAACAACAGCCGCA 793

Qy        302 ACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAACT 361
              |||| | |||||  ||||||| || ||||||| ||| ||| |||| ||||| || |||
Db        794 GCATCCCCTGCATCGCGTGCACGGATGTCAGGAACCCTGTCTTGGTCTTCCAATGTAACC 853

Qy        362 CCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGATC 421
              |||||||||||| || || ||||| |||||||| |||||||| ||||| ||||| ||||
Db        854 ACCGCCACGTGATCTGTTTGGACTGCTTCCACTTGTACTGTGTCACAAGGCTCAACGATC 913

Qy        422 GGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              |||||||||| |||||| |||| |||||||||| ||||| ||||||
Db        914 GGCAGTTTGTCCACGACGCTCAGCTTGGCTACTCGCTGCCGTGTGTG 960


RESULT 14
AF210434
LOCUS       AF210434               1576 bp    mRNA    linear   ROD 02-JUL-2002
DEFINITION  Rattus norvegicus parkin mRNA, complete cds.
ACCESSION   AF210434
VERSION     AF210434.1  GI:11527823
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1576)
  AUTHORS   Gu,W.J., Abbas,N., Lagunes,M.Z., Parent,A., Pradier,L., Bohme,G.A.,
            Agid,Y., Hirsch,E.C., Raisman-Vozari,R. and Brice,A.
  TITLE     Cloning of rat parkin cDNA and distribution of parkin in rat brain
  JOURNAL   J. Neurochem. 74 (4), 1773-1776 (2000)
  MEDLINE   20199898
   PUBMED   10737637
REFERENCE   2  (bases 1 to 1576)
  AUTHORS   Gu,W.J. and Abbas,N.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-DEC-1999) Neurology, INSERM U289, 47, Boulevard de
            L'Hopital, Paris 75013, France
FEATURES             Location/Qualifiers
     source          1. .1576
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /tissue_type="hypothalamus"
     CDS             63. .1460
                     /codon_start=1
                     /product="parkin"
                     /protein_id="AAG37013.1"
                     /db_xref="GI:11527824"
                     /translation="MIVFVRFNSSYGFPVEVDSDTSIFQLKEVVAKRQGVPADQLRVI
                     FAGKELQNHLTVQNCDLEQQSIVHIVQRPQRKSHETNASGGDKPQSTPEGSIWEPRSL
                     TRVDLSSHILPADSVGLAVILDTDSKSDSEAARGPEAKPTYHSFFVYCKGPCHKVQPG
                     KLRVQCGTCRQATLTLAQGPSCWDDVLIPNRMSGECQSPDCPGTRAEFFFKCGAHPTS
                     DKDTSVALNLITNNSRSIPCIACTDVRNPVLVFQCNHRHVICLDCFHLYCVTRLNDRQ
                     FVHDAQLGYSLPCVAGCPNSLIKELHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPR
                     PGCGAGLLPEQGQKKVTCEGGNGLGCGFVFCRDCKEAYHEGECDSMFEASGATSQAYR
                     VDQRAAEQARWEEASKETIKKTTKPCPRCNVPIEKNGGCMHMKCPQPQCKLEWCWNCG
                     CEWNRACMGDHWFDV"
ORIGIN

  Query Match             74.0%;  Score 348.6;  DB 10;  Length 1576;
  Best Local Similarity   84.2%;  Pred. No. 5.4e-103;
  Matches  393;  Conservative    0;  Mismatches   74;  Indels    0;  Gaps    0;

Qy          2 GAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGTC 61
              || |||||| || || | | | |||| ||||||||| ||| || ||||||||||||||| |
Db        466 GAGGTCCAGAAGCTAAACCCACCTACCACAGCTTTTTTGTCTACTGCAAAGGCCCCTGCC 525

Qy         62 AAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTCA 121
              | |  || ||||| || |||||| | || |||||| |||||||||| || ||||| ||||
Db        526 ACAAGGTCCAGCCTGGGAAACTCCGAGTTCAGTGCGGCACCTGCAGACAAGCAACCCTCA 585

Qy        122 CCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGTG 181
              ||||| ||||||| |||||||||||||||||||| ||||||||||||||||||||||| |
Db        586 CCTTGGCCCAGGGCCCATCTTGCTGGGATGATGTCTTAATTCCAAACCGGATGAGTGGAG 645

Qy        182 AATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCAC 241
              | |||||||| ||| |||||||||||| || || ||||||||||||||||||||||||
Db        646 AGTGCCAATCTCCAGACTGCCCTGGGACAAGAGCTGAATTTTTCTTTAAATGTGGAGCAC 705

Qy        242 ACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGGA 301
              |||| ||||| |||||||| |||||||||||||| |||||||| | |  || || || |
Db        706 ACCCAACCTCAGACAAGGACACATCAGTAGCTTTGAACCTGATCACCAACAACAGCCGCA 765

Qy        302 ACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAACT 361
              |||| | |||||  ||||||| || ||||||| ||| ||| |||| ||||| || |||
Db        766 GCATCCCCTGCATCGCGTGCACGGATGTCAGGAACCCTGTCTTGGTCTTCCAATGTAACC 825

Qy        362 CCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGATC 421
              |||||||||||| || || ||||| |||||||| |||||||| ||||| ||||| ||||
Db        826 ACCGCCACGTGATCTGTTTGGACTGCTTCCACTTGTACTGTGTCACAAGGCTCAACGATC 885

Qy        422 GGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              |||||||||| |||||| |||| |||||||||| ||||| ||||||
Db        886 GGCAGTTTGTCCACGACGCTCAGCTTGGCTACTCGCTGCCGTGTGTG 932


RESULT 15
AF343574
LOCUS       AF343574               1157 bp    mRNA    linear   ROD 02-FEB-2002
DEFINITION  Rattus norvegicus parkin transcript variant 6 mRNA, complete cds.
ACCESSION   AF343574
VERSION     AF343574.1  GI:18478865
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1157)
  AUTHORS   D'Agata,V.M., Scapagnini,G. and Cavallaro,S.
  TITLE     Functional and molecular diversity of parkin in the rat brain
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1157)
  AUTHORS   D'Agata,V.M., Scapagnini,G. and Cavallaro,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JAN-2001) Institute of Bioimaging and Pathophysiology
            of the Central Nervous System, Italian National Research Council
            (CNR), Piazza Roma, 2, Catania 95123, Italy
FEATURES             Location/Qualifiers
     source          1. .1157
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
     CDS             333. .1157
                     /codon_start=1
                     /product="parkin transcript variant 6"
                     /protein_id="AAL73348.1"
                     /db_xref="GI:18478866"
                     /translation="MSGECQSPDCPGTRAEFFFKCGAHPTSDKDTSVALNLITNNSRS
                     IPCIACTDVRNPVLVFQCNHRHVICLDCFHLYCVTRLNDRQFVHDAQLGYSLPCVAGC
                     PNSLIKELHHFRILGEEQYNRYQQYGAEECVLQMGGVLCPRPGCGAGLLPEQGQKKVT
                     CEGGNGLGCGFVFCRDCKEAYHEGECDSMFEASGATSQAYRVDQRAAEQARWEEASKE
                     TIKKTTKPCPRCNVPIEKNGGCMHMKCPQPQCKLEWCWNCGCEWNRACMGDHWFDV"
ORIGIN

  Query Match             73.9%;  Score 348.2;  DB 10;  Length 1157;
  Best Local Similarity   84.3%;  Pred. No. 7e-103;
```

Matches 392; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

```
Qy          4 AGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGTCAA 63
              ||| |||||| || | | | |||| ||||||||| ||| || ||||||||||||||| ||
Db        165 AGTACAGCAGCTAAACCCACCTACCACAGCTTTTTTGTCTACTGCAAAGGCCCCTGCCAC 224

Qy         64 AGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTCACC 123
              |  || ||||| || |||||| | || |||||| |||||||||| || ||||| ||||||
Db        225 AAGGTCCAGCCTGGGAAACTCCGAGTTCAGTGCGGCACCTGCAGACAAGCAACCCTCACC 284

Qy        124 TTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGTGAA 183
              ||| ||||||| |||||||||||||||||||| |||||||||||||||||||||||| ||
Db        285 TTGGCCCAGGGCCCATCTTGCTGGGATGATGTCTTAATTCCAAACCGGATGAGTGGAGAG 344

Qy        184 TGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCACAC 243
              |||||||| ||| |||||||||||| || || ||||||||||||||||||||||||||
Db        345 TGCCAATCTCCAGACTGCCCTGGGACAAGAGCTGAATTTTTCTTTAAATGTGGAGCACAC 404

Qy        244 CCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGGAAC 303
              || ||||| |||||||| |||||||||||||| |||||||| | |  || || || | |
Db        405 CCAACCTCAGACAAGGACACATCAGTAGCTTTGAACCTGATCACCAACAACAGCCGCAGC 464

Qy        304 ATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAACTCC 363
              ||| | |||||  ||||||| || ||||||| ||| ||| |||| ||||| || |||  |
Db        465 ATCCCCTGCATCGCGTGCACGGATGTCAGGAACCCTGTCTTGGTCTTCCAATGTAACCAC 524

Qy        364 CGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGATCGG 423
              |||||||||| || || ||||| |||||||| |||||||| ||||| ||||| ||||||
Db        525 CGCCACGTGATCTGTTTGGACTGCTTCCACTTGTACTGTGTCACAAGGCTCAACGATCGG 584

Qy        424 CAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              |||||||| |||||| |||| |||||||||| ||||| ||||||
Db        585 CAGTTTGTCCACGACGCTCAGCTTGGCTACTCGCTGCCGTGTGTG 629
```

Search completed: February 26, 2005, 02:12:59
Job time : 2105.69 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 25, 2005, 15:46:49 ; Search time 265.776 Seconds
(without alignments)
10490.795 Million cell updates/sec

Title: US-10-622-817-10
Perfect score: 471
Sequence: 1 ggaagtccagcaggtagatc.....actccctgccttgtgtgtag 471

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	471	100.0	471	5	AAH77664	Aah77664 Nucleotid
2	468	99.4	2960	3	AAA46765	Aaa46765 cDNA sequ
3	466.4	99.0	2960	2	AAX99923	Aax99923 Human par
4	466.4	99.0	2960	8	AAD47679	Aad47679 Human Par
5	466.4	99.0	2960	8	AAD47680	Aad47680 Human Par
6	466.4	99.0	2960	8	ACD13332	Acd13332 Human DNA
7	340.2	72.2	3043	4	AAF55253	Aaf55253 Nucleotid
8	340.2	72.2	3092	4	AAF55257	Aaf55257 Nucleotid
9	340.2	72.2	3253	4	AAF55256	Aaf55256 Nucleotid
10	340.2	72.2	3253	4	AAF55254	Aaf55254 Nucleotid
11	340.2	72.2	3254	4	AAF55255	Aaf55255 Nucleotid
12	340.2	72.2	3255	4	AAF55259	Aaf55259 Nucleotid
13	340.2	72.2	3255	4	AAF55244	Aaf55244 Nucleotid
14	340.2	72.2	3255	4	AAF55260	Aaf55260 Nucleotid
15	340.2	72.2	3255	4	AAF55258	Aaf55258 Nucleotid
16	338.6	71.9	3014	4	AAF55247	Aaf55247 Nucleotid
17	288.4	61.2	2876	2	AAX99924	Aax99924 Human par
18	288.4	61.2	2876	8	ACD13333	Acd13333 Human DNA
19	265.4	56.3	3136	4	AAF55250	Aaf55250 Nucleotid
20	262.8	55.8	2895	4	AAF55248	Aaf55248 Nucleotid
21	251.8	53.5	2513	8	ACD13334	Acd13334 Human DNA
22	240.8	51.1	857	4	AAF55246	Aaf55246 Nucleotid
23	239.8	50.9	1459	4	AAF55245	Aaf55245 Nucleotid
24	198.8	42.2	3170	4	AAF55251	Aaf55251 Nucleotid
25	100.4	21.3	318	12	ADM96486	Adm96486 Modified
26	90.4	19.2	321	12	ADM96488	Adm96488 Modified
27	78.8	16.7	2918	4	AAF55252	Aaf55252 Nucleotid
28	74	15.7	518	12	ADM96490	Adm96490 Murine pa
29	71	15.1	1458	4	ABL16595	Abl16595 Drosophil
30	53	11.3	470	12	ADM96491	Adm96491 Murine pa
31	50.2	10.7	65	6	ABN53295	Abn53295 Mouse spl
c 32	46.4	9.9	3865	4	ABL16594	Abl16594 Drosophil
33	35.6	7.6	1449	10	ADE07643	Ade07643 Novel cod
34	34.8	7.4	361	3	AAC65989	Aac65989 Human lun
35	34.8	7.4	361	6	ABL49208	Abl49208 Human lun
36	34.8	7.4	361	6	ABQ92394	Abq92394 Human lun
37	34.8	7.4	361	9	ADA28383	Ada28383 Human lun.
38	34.8	7.4	361	10	ADH36947	Adh36947 Human lun
39	34.8	7.4	361	12	ADM56750	Adm56750 Human lun
c 40	34.8	7.4	405	3	AAC07249	Aac07249 Human sec
c 41	33.6	7.1	362	10	ADF80743	Adf80743 Leukaemia
c 42	33.6	7.1	396	6	ABQ56883	Abq56883 Human col
c 43	33.6	7.1	446	12	ADO40914	Ado40914 Human cDN
c 44	33.6	7.1	447	9	ACH26909	Ach26909 Human adu
c 45	33.6	7.1	483	9	ACH26764	Ach26764 Human adu

ALIGNMENTS

RESULT 1
AAH77664
ID AAH77664 standard; DNA; 471 BP.
XX
AC AAH77664;
XX
DT 13-NOV-2001 (first entry)
XX
DE Nucleotide sequence of human Parkin protein fragment.
XX
KW Human; Parkin-Associated Protein 1; PAP1; Parkin gene;
KW neurodegenerative disease; Parkinson's disease; ss.
XX
OS Homo sapiens.
XX
PN WO200160857-A2.
XX
PD 23-AUG-2001.
XX
PF 15-FEB-2001; 2001WO-FR000461.
XX
PR 17-FEB-2000; 2000FR-00001980.
PR 18-APR-2000; 2000US-0198489P.
XX
PA (AVET) AVENTIS PHARMA SA.
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Koutnikova H, Brice A, Fournier A, Pradier L, Prades C;
PI Arnould-Reguigne I, Rosier-Montus M, Corti O;
XX
DR WPI; 2001-550047/61.
DR P-PSDB; AAG67213.
XX
PT A new protein, designated Parkin-Associated Protein 1 (PAP1), is an
PT interaction partner of Parkin and is useful to treat neurodegenerative
PT pathologies including Parkinson's disease.
XX
PS Example 1; Page 60-61; 82pp; French.
XX
CC The present sequence encodes the central region of a human Parkin
CC protein. The protein was used to identify Parkin-Associated Protein 1
CC (PAP1) protein. PAP1 is associated with the Parkin gene, which is mutated
CC in certain forms of familial (juvenile autosomal recessive) Parkinson's

```
CC   disease. PAP1 has some homology with synaptotagmins. PAP1 is used to
CC   treat neurodegenerative diseases, particularly to diagnose and treat
CC   Parkinson's disease
XX
SQ   Sequence 471 BP; 118 A; 127 C; 109 G; 117 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 471;  DB 5;  Length 471;
  Best Local Similarity   100.0%;  Pred. No. 7.8e-153;
  Matches  471;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60

Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120

Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180

Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240

Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300

Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360

Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420

Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTGTAG 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTGTAG 471


RESULT 2
AAA46765
ID   AAA46765 standard; cDNA; 2960 BP.
XX
AC   AAA46765;
XX
DT   25-SEP-2000  (first entry)
XX
DE   cDNA sequence of the human parkin protein gene.
XX
KW   Human; parkin protein; parkin gene; Parkinson's disease;
KW   anti-Parkinson agent; ss.
XX
OS   Homo sapiens.
XX
PN   WO200031253-A2.
XX
PD   02-JUN-2000.
XX
PF   18-NOV-1999;   99WO-FR002833.
XX
PR   19-NOV-1998;   98FR-00014524.
PR   12-MAR-1999;   99US-0124239P.
PR   04-AUG-1999;   99FR-00010140.
XX
PA   (RHON ) RHONE-POULENC RORER SA.
PA   (INRM ) INST NAT SANTE & RECH MEDICALE.
XX
PI   Brice A,  Lucking C,  Abbas NE,  Denefle P,  Ricard S,  Bouley S;
XX
DR   WPI; 2000-411952/35.
XX
PT   New variant forms of the human parkin gene, used as source of primers and
PT   probes for detecting susceptibility to Parkinson's disease.
XX
PS   Claim 24; Fig 1; 71pp; French.
XX
CC   The present sequence represents a cDNA sequence of the human parkin
CC   protein gene. The specification describes a parkin gene which has genetic
CC   alterations. Cells, or transgenic animals, that express the altered
CC   parkin gene are used to screen for compounds that can counter the effects
CC   of a genetic alteration in the parkin gene, or more generally for
CC   studying the properties of the parkin protein. Detection of the specified
CC   alterations is used to diagnose susceptibility to Parkinson's disease.
CC   The modified polynucleotide is also used to express the corresponding
CC   protein, which is then used to screen for potential anti-Parkinson agents
CC   and to raise antibodies (for detecting variants of parkin protein)
XX
SQ   Sequence 2960 BP; 815 A; 726 C; 698 G; 721 T; 0 U; 0 Other;

  Query Match             99.4%;  Score 468;  DB 3;  Length 2960;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-151;
  Matches  468;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 563

Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 623

Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 683

Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 743

Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        744 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 803

Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 863

Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 923

Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 971


RESULT 3
AAX99923
ID   AAX99923 standard; DNA; 2960 BP.
XX
AC   AAX99923;
XX
DT   21-OCT-1999  (first entry)
XX
DE   Human parkin gene variant.
XX
KW   Parkinson's disease related gene; parkin gene; variant; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             102. .1499
```

```
FT                   /*tag=  a
XX
PN   WO9940191-A1.
XX
PD   12-AUG-1999.
XX
PF   09-FEB-1999;   99WO-JP000545.
XX
PR   09-FEB-1998;   98JP-00027531.
XX
PA   (SHIM/) SHIMIZU N.
PA   (MIZU/) MIZUNO Y.
XX
PI   Shimizu N,  Mizuno Y;
XX
DR   WPI; 1999-494295/41.
DR   P-PSDB; AAY32501.
XX
PT   Gene implicated in the pathology of Parkinson's disease, used for
PT   treatment of the disease.
XX
PS   Claim 1; Page 83-88; 114pp; English.
XX
CC   This sequence represents a gene of the invention, and is implicated in
CC   the pathology of Parkinson's disease. This sequence is a variant of the
CC   parkin gene found in parkinson's disease patients. The sequences may be
CC   used for the diagnosis, treatment (including gene therapy) and
CC   investigation of Parkinson's disease
XX
SQ   Sequence 2960 BP; 815 A; 727 C; 698 G; 720 T; 0 U; 0 Other;

  Query Match             99.0%;  Score 466.4;  DB 2;  Length 2960;
  Best Local Similarity   99.8%;  Pred. No. 8.4e-151;
  Matches  467;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 563

Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 623

Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 683

Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 743

Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
              ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        744 CACCCCACCTCTGACAAGGAAACACCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 803

Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 863

Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 923

Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 971


RESULT 4
AAD47679
ID   AAD47679 standard; cDNA; 2960 BP.
XX
AC   AAD47679;
XX
DT   24-FEB-2003  (first entry)
XX
DE   Human Parkin protein encoding cDNA.
XX
KW   Human; Parkin protein; neurological disorder; apoptosis; gene therapy;
KW   ischaemic stroke; Parkinson's disease; Alzheimer's disease; nootropic;
KW   transgenic; cerebroprotective; neuroprotective; neurotransplantation;
KW   gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             102. .1499
FT                   /*tag=  a
FT                   /product= "Human Parkin protein"
XX
PN   WO200279459-A2.
XX
PD   10-OCT-2002.
XX
PF   02-APR-2002; 2002WO-DK000221.
XX
PR   29-MAR-2001; 2001DK-00000525.
PR   03-APR-2001; 2001US-0281286P.
XX
PA   (NSGE-) NSGENE AS.
XX
PI   Jensen PH;
XX
DR   WPI; 2003-046812/04.
DR   P-PSDB; AAE30800.
XX
PT   New isolated nucleic acid sequence encoding a Parkin polypeptide, useful
PT   for treating, preventing or diagnosing neurological disorders, e.g.
PT   Parkinson's disease, Alzheimer's disease or ischemic stroke, and in
PT   screening assays.
XX
PS   Example 1; Page 68-69; 71pp; English.
XX
CC   The invention relates to Parkin protein and its corresponding nucleic
CC   acid sequence. The nucleic acid sequence is useful for altering the
CC   proteolytic processing of Parkin at its potential cleavage site at Asp
CC   126. The invention is used in manufacturing or testing a pharmaceutical
CC   composition for treating and/or preventing a neurological disorder, e.g.
CC   Alzheimer's disease or ischaemic stroke. It also used for detecting the
CC   occurrence of proteolytic processing of Parkin at Asp 126 in a sample, in
CC   monitoring a potential disposition for a neurodegenerative disease, and
CC   for treating, preventing and/or diagnosing Parkinson's disease or other
CC   neurodegenerative disorders. The viral vector is used for transforming
CC   neuronal cells in vivo or ex vivo. The invention is useful for
CC   neurotransplantation into the CNS of a mammal. It may be used in
CC   screening assays to identify compounds that increase or decrease
CC   apoptosis. It is also used in gene therapy. The present sequence is human
CC   Parkin protein encoding cDNA
XX
SQ   Sequence 2960 BP; 815 A; 727 C; 698 G; 720 T; 0 U; 0 Other;

  Query Match             99.0%;  Score 466.4;  DB 8;  Length 2960;
  Best Local Similarity   99.8%;  Pred. No. 8.4e-151;
  Matches  467;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 563

Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 623

Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 683
```

```
Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 743

Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        744 CACCCCACCTCTGACAAGGAAACACCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 803

Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 863

Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 923

Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        924 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 971
```

RESULT 5
AAD47680
ID AAD47680 standard; DNA; 2960 BP.
XX
AC AAD47680;
XX
DT 24-FEB-2003 (first entry)
XX
DE Human Parkin D126 mutant DNA.
XX
KW Human; Parkin protein; neurological disorder; apoptosis; gene therapy;
KW ischaemic stroke; Parkinson's disease; Alzheimer's disease; nootropic;
KW transgenic; cerebroprotective; neuroprotective; neurotransplantation;
KW mutant; gene; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 102. .1499
FT /*tag= a
FT /product= "Human Parkin D126 mutant protein"
FT /transl_except= (pos:768. .770, aa:Ser)
XX
PN WO200279459-A2.
XX
PD 10-OCT-2002.
XX
PF 02-APR-2002; 2002WO-DK000221.
XX
PR 29-MAR-2001; 2001DK-00000525.
PR 03-APR-2001; 2001US-0281286P.
XX
PA (NSGE-) NSGENE AS.
XX
PI Jensen PH;
XX
DR WPI; 2003-046812/04.
DR P-PSDB; AAE30801.
XX
PT New isolated nucleic acid sequence encoding a Parkin polypeptide, useful
PT for treating, preventing or diagnosing neurological disorders, e.g.
PT Parkinson's disease, Alzheimer's disease or ischemic stroke, and in
PT screening assays.
XX
PS Claim 4; Page 69-71; 71pp; English.
XX
CC The invention relates to Parkin protein and its corresponding nucleic
CC acid sequence. The nucleic acid sequence is useful for altering the
CC proteolytic processing of Parkin at its potential cleavage site at Asp
CC 126. The invention is used in manufacturing or testing a pharmaceutical
CC composition for treating and/or preventing a neurological disorder, e.g.
CC Alzheimer's disease or ischaemic stroke. It also used for detecting the
CC occurrence of proteolytic processing of Parkin at Asp 126 in a sample, in
CC monitoring a potential disposition for a neurodegenerative disease, and
CC for treating, preventing and/or diagnosing Parkinson's disease or other
CC neurodegenerative disorders. The viral vector is used for transforming
CC neuronal cells in vivo or ex vivo. The invention is useful for
CC neurotransplantation into the CNS of a mammal. It may be used in
CC screening assays to identify compounds that increase or decrease
CC apoptosis. It is also used in gene therapy. The present sequence is human
CC Parkin D126 mutant DNA
XX
SQ Sequence 2960 BP; 816 A; 726 C; 698 G; 720 T; 0 U; 0 Other;

Query Match 99.0%; Score 466.4; DB 8; Length 2960;
Best Local Similarity 99.8%; Pred. No. 8.4e-151;
Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 563

Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 623

Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 683

Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 743

Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        744 CACCCCACCTCTGACAAGGAAACACCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 803

Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 863

Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 923

Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        924 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 971
```

RESULT 6
ACD13332
ID ACD13332 standard; cDNA; 2960 BP.
XX
AC ACD13332;
XX
DT 13-AUG-2003 (first entry)
XX
DE Human DNA encoding a p53 modifier, SEQ ID 1.
XX
KW Human; ss; gene; p53 modifier; cytostatic; cancer; cytostatic;
KW antiangiogenic; antiapoptotic; p53 pathway; breast cancer; colon cancer;
KW kidney cancer; lung cancer; ovarian cancer; angiogenesis; cell cycle;
KW apoptotic disorder; cell proliferation disorder.
XX
OS Homo sapiens.
XX
PN WO200299122-A1.
XX
PD 12-DEC-2002.
XX
PF 03-JUN-2002; 2002WO-US017382.

```
XX
PR   05-JUN-2001; 2001US-0296076P.
PR   10-OCT-2001; 2001US-0328605P.
PR   15-FEB-2002; 2002US-0357253P.
XX
PA   (EXEL-) EXELIXIS INC.
XX
PI   Friedman L,  Plowman GD,  Belvin M,  Francis-Lang H,  Li D,  Funke RP;
XX
DR   WPI; 2003-156859/15.
DR   P-PSDB; ABO07156.
XX
PT   Identifying modulators of the p53 pathway for use in treating apoptotic
PT   or cell proliferation disorders, comprises screening for agents that
PT   modulate activity of a human ortholog of genes that modify the p53
PT   pathway in Drosophila.
XX
PS   Example 2; Page 166-167; 678pp; English.
XX
CC   The invention relates to identifying (M1) a candidate p53 pathway
CC   modulating agent, by contacting an assay system comprising a purified HM
CC   polypeptide (human orthologue of genes that modify the p53 pathway in
CC   Drosophila) or nucleic acid with a test agent under conditions, where but
CC   for the presence of the test agent, the system provides a reference
CC   activity, and detecting a test agent-biased activity of the assay system.
CC   Also included are modulating (M2) a p53 pathway of a cell (comprising
CC   contacting a cell defective in p53 function with a candidate modulator
CC   that specifically binds to a HM polypeptide comprising an HM amino acid
CC   sequence, where p53 function is restored), modulating (M3) a p53 pathway
CC   in a mammalian cell (comprising contacting the cell with an agent that
CC   specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)
CC   a disease in a patient (comprising: (a) obtaining a biological sample
CC   from the patient; (b) contacting the sample with a probe for HM
CC   expression; (c) comparing the results with a control; and (d) determining
CC   whether the comparison indicates a likelihood disease). (M1) is useful
CC   for identifying modulators of the p53 pathway. A probe for HM expression
CC   is useful for diagnosing breast, colon, kidney, lung and ovarian cancer,
CC   in a patient, where the cancer has greater than 25 % expression level.
CC   Modulators identified by (M1) are useful in a variety of diagnostic and
CC   therapeutic applications, where disease or disorder prognosis is related
CC   to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell
CC   proliferation disorders (e.g. cancer). Another two new methods (M2 and
CC   M3) are useful for modulating the p53 pathway of a cell, thus restoring
CC   the p53 function of the cell, so that the cell undergoes normal
CC   proliferation or progression through the cell cycle. (M2) and (M3) are
CC   also useful for treating defects in the p53 pathway such as angiogenic,
CC   apoptotic or cell proliferation disorders. The present sequence is an HM
CC   nucleic acid encoding a p53 pathway modifying protein
XX
SQ   Sequence 2960 BP; 815 A; 727 C; 698 G; 720 T; 0 U; 0 Other;

  Query Match           99.0%;  Score 466.4;  DB 8;  Length 2960;
  Best Local Similarity  99.8%;  Pred. No. 8.4e-151;
  Matches  467;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 563

Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 623

Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 683

Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 743

Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        744 CACCCCACCTCTGACAAGGAAACACCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 803

Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 863

Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 923

Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 971


RESULT 7
AAF55253
ID   AAF55253 standard; cDNA; 3043 BP.
XX
AC   AAF55253;
XX
DT   29-MAY-2001  (first entry)
XX
DE   Nucleotide sequence of murine parkin2 with a frameshift mutation.
XX
KW   Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW   Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW   Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW   brain tumour; head trauma; stroke; vascular irregularity;
KW   metabolic irregularity; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             129. .1019
FT                   /*tag=  a
FT                   /product= "truncated parkin2"
XX
PN   EP1081225-A1.
XX
PD   07-MAR-2001.
XX
PF   30-AUG-1999;   99EP-00116766.
XX
PR   30-AUG-1999;   99EP-00116766.
XX
PA   (BIOF-) BIOFRONTERA PHARM GMBH.
XX
PI   Luebbert H;
XX
DR   WPI; 2001-212797/22.
DR   P-PSDB; AAB67526.
XX
PT   New polynucleotides encoding mouse parkin2 protein, useful for producing
PT   a transgenic non-human animal as an animal model for neurodegenerative
PT   diseases.
XX
PS   Claim 3; Page 29-30; 62pp; English.
XX
CC   The present sequence encodes a murine parkin2 polypeptide. The
CC   polynucleotide sequence contains a frameshift mutation, leading to a
CC   truncated protein. Mutations or deletions in the parkin2 gene cause
CC   Parkinson's disease in humans. The human parkin2 gene is located in gene
CC   region 6q25.2-27. Parkin2 polypeptides and polynucleotides are useful for
CC   analysing neurodegenerative diseases. They are also useful for testing
CC   the efficacy of the treatment of a neurodegenerative disease such as
CC   Parkinson's disease, Alzheimer's disease, Huntington's disease,
CC   amyotrophic lateral sclerosis, Multi-system atrophy, Wilson's disease,
CC   Pick's disease, Prion disease, and secondary causes inducing Parkinson's
CC   syndromes like toxins, drugs, brain tumours, head trauma, stroke,
CC   vascular irregularities or metabolic irregularities, associated with a
CC   less active or non-active parkin protein
XX
```

```
SQ   Sequence 3043 BP; 806 A; 763 C; 737 G; 737 T; 0 U; 0 Other;

  Query Match             72.2%;  Score 340.2;  DB 4;  Length 3043;
  Best Local Similarity   84.0%;  Pred. No. 7e-107;
  Matches  384;  Conservative    0;  Mismatches   73;  Indels    0;  Gaps    0;

Qy          11 CAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGTCAAAGAGTGC 70
               ||| || | | | ||||||||||||||    | || ||||||||||||||| || |  || |
Db         538 CAGTTAAACCCACCTACAACAGCTTTTTCATCTACTGCAAAGGCCCCTGCCACAAGGTCC 597

Qy          71 AGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTCACCTTGACCC 130
               |||| ||||| ||| | || |||||  |||||||||  || ||||| ||||||||| |||
Db         598 AGCCTGGAAAGCTCCGAGTTCAGTGTGGCACCTGCAAACAAGCAACCCTCACCTTGGCCC 657

Qy         131 AGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGTGAATGCCAAT 190
               |||| |||||||||||||| ||||| |||||||||||||||||||||||||| ||||| |
Db         658 AGGGCCCATCTTGCTGGGACGATGTCTTAATTCCAAACCGGATGAGTGGTGAGTGCCAGT 717

Qy         191 CCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCACACCCCACCT 250
               | ||| |||||||||| || || || |||||||||||||||||||||||||||| ||||
Db         718 CTCCAGACTGCCCTGGAACCAGAGCTGAATTTTTCTTTAAATGTGGAGCACACCCAACCT 777

Qy         251 CTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGGAACATCACTT 310
               | |||||||| || || ||||||||| |||||||| | |  || || || | |||| |||
Db         778 CAGACAAGGACACGTCGGTAGCTTTGAACCTGATCACCAGCAACAGGCGCAGCATCCCTT 837

Qy         311 GCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAACTCCCGCCACG 370
               ||||  |||||||||| |||||||||| |||||||| |||||||| |||  ||| ||||
Db         838 GCATAGCGTGCACAGATGTCAGGAGCCCTGTCCTGGTCTTCCAGTGTAACCACCGTCACG 897

Qy         371 TGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGATCGGCAGTTTG 430
               |||| || || |||||||||||||| || ||||| |||||||||| ||||||||||||
Db         898 TGATCTGTTTGGACTGTTTCCACTTGTATTGTGTCACAAGACTCAACGATCGGCAGTTTG 957

Qy         431 TTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGT 467
               | |||||  |||||||||||||||||||||| |||||
Db         958 TCCACGATGCTCAACTTGGCTACTCCCTGCCGTGTGT 994


RESULT 8
AAF55257
ID   AAF55257 standard; cDNA; 3092 BP.
XX
AC   AAF55257;
XX
DT   29-MAY-2001  (first entry)
XX
DE   Nucleotide sequence of murine parkin2 with a frameshift mutation.
XX
KW   Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW   Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW   Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW   brain tumour; head trauma; stroke; vascular irregularity;
KW   metabolic irregularity; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             129. .173
FT                   /*tag= a
FT                   /product= "truncated parkin2"
XX
PN   EP1081225-A1.
XX
PD   07-MAR-2001.
XX
PF   30-AUG-1999;   99EP-00116766.
XX
PR   30-AUG-1999;   99EP-00116766.
XX
PA   (BIOF-) BIOFRONTERA PHARM GMBH.
XX
PI   Luebbert H;
XX
DR   WPI; 2001-212797/22.
DR   P-PSDB; AAB67530.
XX
PT   New polynucleotides encoding mouse parkin2 protein, useful for producing
PT   a transgenic non-human animal as an animal model for neurodegenerative
PT   diseases.
XX
PS   Claim 3; Page 34-35; 62pp; English.
XX
CC   The present sequence encodes a murine parkin2 polypeptide. The
CC   polynucleotide sequence contains a frameshift mutation, leading to a
CC   truncated protein. Mutations or deletions in the parkin2 gene cause
CC   Parkinson's disease in humans. The human parkin2 gene is located in gene
CC   region 6q25.2-27. Parkin2 polypeptides and polynucleotides are useful for
CC   analysing neurodegenerative diseases. They are also useful for testing
CC   the efficacy of the treatment of a neurodegenerative disease such as
CC   Parkinson's disease, Alzheimer's disease, Huntington's disease,
CC   amyotrophic lateral sclerosis, Multi-system atrophy, Wilson's disease,
CC   Pick's disease, Prion disease, and secondary causes inducing Parkinson's
CC   syndromes like toxins, drugs, brain tumours, head trauma, stroke,
CC   vascular irregularities or metabolic irregularities, associated with a
CC   less active or non-active parkin protein
XX
SQ   Sequence 3092 BP; 818 A; 778 C; 761 G; 735 T; 0 U; 0 Other;

  Query Match             72.2%;  Score 340.2;  DB 4;  Length 3092;
  Best Local Similarity   84.0%;  Pred. No. 7e-107;
  Matches  384;  Conservative    0;  Mismatches   73;  Indels    0;  Gaps    0;

Qy          11 CAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGTCAAAGAGTGC 70
               ||| || | | | ||||||||||||||    | || ||||||||||||||| || |  || |
Db         375 CAGTTAAACCCACCTACAACAGCTTTTTCATCTACTGCAAAGGCCCCTGCCACAAGGTCC 434

Qy          71 AGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTCACCTTGACCC 130
               |||| ||||| ||| | || |||||  |||||||||  || ||||| ||||||||| |||
Db         435 AGCCTGGAAAGCTCCGAGTTCAGTGTGGCACCTGCAAACAAGCAACCCTCACCTTGGCCC 494

Qy         131 AGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGTGAATGCCAAT 190
               |||| |||||||||||||| ||||| |||||||||||||||||||||||||| ||||| |
Db         495 AGGGCCCATCTTGCTGGGACGATGTCTTAATTCCAAACCGGATGAGTGGTGAGTGCCAGT 554

Qy         191 CCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCACACCCCACCT 250
               | ||| |||||||||| || || || |||||||||||||||||||||||||||| ||||
Db         555 CTCCAGACTGCCCTGGAACCAGAGCTGAATTTTTCTTTAAATGTGGAGCACACCCAACCT 614

Qy         251 CTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGGAACATCACTT 310
               | |||||||| || || ||||||||| |||||||| | |  || || || | |||| |||
Db         615 CAGACAAGGACACGTCGGTAGCTTTGAACCTGATCACCAGCAACAGGCGCAGCATCCCTT 674

Qy         311 GCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAACTCCCGCCACG 370
               ||||  |||||||||| |||||||||| |||||||| |||||||| |||  ||| ||||
Db         675 GCATAGCGTGCACAGATGTCAGGAGCCCTGTCCTGGTCTTCCAGTGTAACCACCGTCACG 734

Qy         371 TGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGATCGGCAGTTTG 430
               |||| || || |||||||||||||| || ||||| |||||||||| ||||||||||||
Db         735 TGATCTGTTTGGACTGTTTCCACTTGTATTGTGTCACAAGACTCAACGATCGGCAGTTTG 794

Qy         431 TTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGT 467
               | |||||  |||||||||||||||||||||| |||||
Db         795 TCCACGATGCTCAACTTGGCTACTCCCTGCCGTGTGT 831


RESULT 9
AAF55256
ID   AAF55256 standard; cDNA; 3253 BP.
XX
AC   AAF55256;
XX
DT   29-MAY-2001  (first entry)
XX
DE   Nucleotide sequence of murine parkin2 with a frameshift mutation.
XX
```

```
KW   Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW   Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW   Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW   brain tumour; head trauma; stroke; vascular irregularity;
KW   metabolic irregularity; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             129. .362
PT                   /*tag= a
FT                   /product= "truncated parkin2"
XX
PN   EP1081225-A1.
XX
PD   07-MAR-2001.
XX
PF   30-AUG-1999;   99EP-00116766.
XX
PR   30-AUG-1999;   99EP-00116766.
XX
PA   (BIOF-) BIOFRONTERA PHARM GMBH.
XX
PI   Luebbert H;
XX
DR   WPI; 2001-212797/22.
DR   P-PSDB; AAB67529.
XX
PT   New polynucleotides encoding mouse parkin2 protein, useful for producing
PT   a transgenic non-human animal as an animal model for neurodegenerative
PT   diseases.
XX
PS   Claim 3; Page 33-34; 62pp; English.
XX
CC   The present sequence encodes a murine parkin2 polypeptide. The
CC   polynucleotide sequence contains mutations, causing Arg78 to be changed
CC   to a stop codon, leading to a truncated protein. Mutations or deletions
CC   in the parkin2 gene cause Parkinson's disease in humans. The human
CC   parkin2 gene is located in gene region 6q25.2-27. Parkin2 polypeptides
CC   and polynucleotides are useful for analysing neurodegenerative diseases.
CC   They are also useful for testing the efficacy of the treatment of a
CC   neurodegenerative disease such as Parkinson's disease, Alzheimer's
CC   disease, Huntington's disease, amyotrophic lateral sclerosis, Multi-
CC   system atrophy, Wilson's disease, Pick's disease, Prion disease, and
CC   secondary causes inducing Parkinson's syndromes like toxins, drugs, brain
CC   tumours, head trauma, stroke, vascular irregularities or metabolic
CC   irregularities, associated with a less active or non-active parkin
CC   protein
XX
SQ   Sequence 3253 BP; 851 A; 818 C; 808 G; 776 T; 0 U; 0 Other;

  Query Match           72.2%;  Score 340.2;  DB 4;  Length 3253;
  Best Local Similarity  84.0%;  Pred. No. 7.2e-107;
  Matches  384;  Conservative    0;  Mismatches   73;  Indels    0;  Gaps    0;

Qy         11 CAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGTCAAAGAGTGC 70
              ||| || | | |||||||||||||||   | || ||||||||||||||| || |  || |
Db        536 CAGTTAAACCCACCTACAACAGCTTTTTCATCTACTGCAAAGGCCCCTGCCACAAGGTCC 595

Qy         71 AGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTCACCTTGACCC 130
              |||| ||||| ||| | || |||||  |||||||||  || |||||  ||||||||| |||
Db        596 AGCCTGGAAAGCTCCGAGTTCAGTGTGGCACCTGCAAACAAGCAACCCTCACCTTGGCCC 655

Qy        131 AGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGTGAATGCCAAT 190
              |||| |||||||||||||| ||||| |||||||||||||||||||||||||||| ||||| |
Db        656 AGGGCCCATCTTGCTGGGACGATGTCTTAATTCCAAACCGGATGAGTGGTGAGTGCCAGT 715

Qy        191 CCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCACACCCCACCT 250
              | ||| |||||||||| || || || |||||||||||||||||||||||||||||| ||||
Db        716 CTCCAGACTGCCCTGGAACCAGAGCTGAATTTTTCTTTAAATGTGGAGCACACCCAACCT 775

Qy        251 CTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGGAACATCACTT 310
              | |||||||| || || ||||||||| |||||||| | |  || || || | |||| |||
Db        776 CAGACAAGGACACGTCGGTAGCTTTGAACCTGATCACCAGCAACAGGCGCAGCATCCCTT 835

Qy        311 GCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAACTCCCGCCACG 370
              ||||  |||||||||| |||||||||| |||||||| |||||||| |||  ||| ||||
Db        836 GCATAGCGTGCACAGATGTCAGGAGCCCTGTCCTGGTCTTCCAGTGTAACCACCGTCACG 895

Qy        371 TGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGATCGGCAGTTTG 430
              |||| || || |||||||||||||| || ||||| |||||||||| ||||||||||||
Db        896 TGATCTGTTTGGACTGTTTCCACTTGTATTGTGTCACAAGACTCAACGATCGGCAGTTTG 955

Qy        431 TTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGT 467
              | |||||  |||||||||||||||||||||| |||||
Db        956 TCCACGATGCTCAACTTGGCTACTCCCTGCCGTGTGT 992


RESULT 10
AAF55254
ID   AAF55254 standard; cDNA; 3253 BP.
XX
AC   AAF55254;
XX
DT   29-MAY-2001  (first entry)
XX
DE   Nucleotide sequence of murine parkin2 with a frameshift mutation.
XX
KW   Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW   Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW   Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW   brain tumour; head trauma; stroke; vascular irregularity;
KW   metabolic irregularity; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             129. .242
FT                   /*tag= a
FT                   /product= "truncated parkin2"
XX
PN   EP1081225-A1.
XX
PD   07-MAR-2001.
XX
PF   30-AUG-1999;   99EP-00116766.
XX
PR   30-AUG-1999;   99EP-00116766.
XX
PA   (BIOF-) BIOFRONTERA PHARM GMBH.
XX
PI   Luebbert H;
XX
DR   WPI; 2001-212797/22.
DR   P-PSDB; AAB67527.
XX
PT   New polynucleotides encoding mouse parkin2 protein, useful for producing
PT   a transgenic non-human animal as an animal model for neurodegenerative
PT   diseases.
XX
PS   Claim 3; Page 30-31; 62pp; English.
XX
CC   The present sequence encodes a murine parkin2 polypeptide. The
CC   polynucleotide sequence contains mutations, causing Gln38 to be changed
CC   to a stop codon, leading to a truncated protein. Mutations or deletions
CC   in the parkin2 gene cause Parkinson's disease in humans. The human
CC   parkin2 gene is located in gene region 6q25.2-27. Parkin2 polypeptides
CC   and polynucleotides are useful for analysing neurodegenerative diseases.
CC   They are also useful for testing the efficacy of the treatment of a
CC   neurodegenerative disease such as Parkinson's disease, Alzheimer's
CC   disease, Huntington's disease, amyotrophic lateral sclerosis, Multi-
CC   system atrophy, Wilson's disease, Pick's disease, Prion disease, and
CC   secondary causes inducing Parkinson's syndromes like toxins, drugs, brain
CC   tumours, head trauma, stroke, vascular irregularities or metabolic
CC   irregularities, associated with a less active or non-active parkin
CC   protein
```

```
XX
SQ   Sequence 3253 BP; 851 A; 818 C; 808 G; 776 T; 0 U; 0 Other;

  Query Match            72.2%;  Score 340.2;  DB 4;  Length 3253;
  Best Local Similarity   84.0%;  Pred. No. 7.2e-107;
  Matches  384;  Conservative    0;  Mismatches   73;  Indels    0;  Gaps    0;

Qy         11 CAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGTCAAAGAGTGC 70
              ||| || | | | ||||||||||||||    | || ||||||||||||| || |  || |
Db        536 CAGTTAAACCCACCTACAACAGCTTTTTCATCTACTGCAAAGGCCCCTGCCACAAGGTCC 595

Qy         71 AGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTCACCTTGACCC 130
              |||| ||||| ||| | || |||||  |||||||||  || ||||| ||||||||| |||
Db        596 AGCCTGGAAAGCTCCGAGTTCAGTGTGGCACCTGCAAACAAGCAACCCTCACCTTGGCCC 655

Qy        131 AGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGTGAATGCCAAT 190
              |||| ||||||||||||| ||||| |||||||||||||||||||||||||||||| |||||
Db        656 AGGGCCCATCTTGCTGGGACGATGTCTTAATTCCAAACCGGATGAGTGGTGAGTGCCAGT 715

Qy        191 CCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCACACCCCACCT 250
              | ||| |||||||||| || || || |||||||||||||||||||||||||||||| ||||
Db        716 CTCCAGACTGCCCTGGAACCAGAGCTGAATTTTTCTTTAAATGTGGAGCACACCCAACCT 775

Qy        251 CTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGGAACATCACTT 310
              | |||||||| || || ||||||||| |||||||| | |  || || || | |||| |||
Db        776 CAGACAAGGACACGTCGGTAGCTTTGAACCTGATCACCAGCAACAGGCGCAGCATCCCTT 835

Qy        311 GCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAACTCCCGCCACG 370
              ||||  |||||||||| |||||||||| |||||||| |||||||| |||  ||| ||||
Db        836 GCATAGCGTGCACAGATGTCAGGAGCCCTGTCCTGGTCTTCCAGTGTAACCACCGTCACG 895

Qy        371 TGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGATCGGCAGTTTG 430
              |||| || || |||||||||||||| || ||||| |||||||||| ||||||||||||
Db        896 TGATCTGTTTGGACTGTTTCCACTTGTATTGTGTCACAAGACTCAACGATCGGCAGTTTG 955

Qy        431 TTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGT 467
              | |||||  |||||||||||||||||||||| |||||
Db        956 TCCACGATGCTCAACTTGGCTACTCCCTGCCGTGTGT 992


RESULT 11
AAF55255
ID   AAF55255 standard; cDNA; 3254 BP.
XX
AC   AAF55255;
XX
DT   29-MAY-2001  (first entry)
XX
DE   Nucleotide sequence of murine parkin2 with a frameshift mutation.
XX
KW   Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW   Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW   Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW   brain tumour; head trauma; stroke; vascular irregularity;
KW   metabolic irregularity; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             129. .290
FT                   /*tag=  a
FT                   /product= "truncated parkin2"
XX
PN   EP1081225-A1.
XX
PD   07-MAR-2001.
XX
PF   30-AUG-1999;   99EP-00116766.
XX
PR   30-AUG-1999;   99EP-00116766.
XX
PA   (BIOF-) BIOFRONTERA PHARM GMBH.
XX
PI   Luebbert H;
XX
DR   WPI; 2001-212797/22.
DR   P-PSDB; AAB67528.
XX
PT   New polynucleotides encoding mouse parkin2 protein, useful for producing
PT   a transgenic non-human animal as an animal model for neurodegenerative
PT   diseases.
XX
PS   Claim 3; Page 31-32; 62pp; English.
XX
CC   The present sequence encodes a murine parkin2 polypeptide. The
CC   polynucleotide sequence contains mutations, causing Asn54 to be changed
CC   to a stop codon, leading to a truncated protein. Mutations or deletions
CC   in the parkin2 gene cause Parkinson's disease in humans. The human
CC   parkin2 gene is located in gene region 6q25.2-27. Parkin2 polypeptides
CC   and polynucleotides are useful for analysing neurodegenerative diseases.
CC   They are also useful for testing the efficacy of the treatment of a
CC   neurodegenerative disease such as Parkinson's disease, Alzheimer's
CC   disease, Huntington's disease, amyotrophic lateral sclerosis, Multi-
CC   system atrophy, Wilson's disease, Pick's disease, Prion disease, and
CC   secondary causes inducing Parkinson's syndromes like toxins, drugs, brain
CC   tumours, head trauma, stroke, vascular irregularities or metabolic
CC   irregularities, associated with a less active or non-active parkin
CC   protein
XX
SQ   Sequence 3254 BP; 851 A; 818 C; 809 G; 776 T; 0 U; 0 Other;

  Query Match            72.2%;  Score 340.2;  DB 4;  Length 3254;
  Best Local Similarity   84.0%;  Pred. No. 7.2e-107;
  Matches  384;  Conservative    0;  Mismatches   73;  Indels    0;  Gaps    0;

Qy         11 CAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGTCAAAGAGTGC 70
              ||| || | | | ||||||||||||||    | || ||||||||||||| || |  || |
Db        537 CAGTTAAACCCACCTACAACAGCTTTTTCATCTACTGCAAAGGCCCCTGCCACAAGGTCC 596

Qy         71 AGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTCACCTTGACCC 130
              |||| ||||| ||| | || |||||  |||||||||  || ||||| ||||||||| |||
Db        597 AGCCTGGAAAGCTCCGAGTTCAGTGTGGCACCTGCAAACAAGCAACCCTCACCTTGGCCC 656

Qy        131 AGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGTGAATGCCAAT 190
              |||| ||||||||||||| ||||| |||||||||||||||||||||||||||||| |||||
Db        657 AGGGCCCATCTTGCTGGGACGATGTCTTAATTCCAAACCGGATGAGTGGTGAGTGCCAGT 716

Qy        191 CCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCACACCCCACCT 250
              | ||| |||||||||| || || || |||||||||||||||||||||||||||||| ||||
Db        717 CTCCAGACTGCCCTGGAACCAGAGCTGAATTTTTCTTTAAATGTGGAGCACACCCAACCT 776

Qy        251 CTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGGAACATCACTT 310
              | |||||||| || || ||||||||| |||||||| | |  || || || | |||| |||
Db        777 CAGACAAGGACACGTCGGTAGCTTTGAACCTGATCACCAGCAACAGGCGCAGCATCCCTT 836

Qy        311 GCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAACTCCCGCCACG 370
              ||||  |||||||||| |||||||||| |||||||| |||||||| |||  ||| ||||
Db        837 GCATAGCGTGCACAGATGTCAGGAGCCCTGTCCTGGTCTTCCAGTGTAACCACCGTCACG 896

Qy        371 TGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGATCGGCAGTTTG 430
              |||| || || |||||||||||||| || ||||| |||||||||| ||||||||||||
Db        897 TGATCTGTTTGGACTGTTTCCACTTGTATTGTGTCACAAGACTCAACGATCGGCAGTTTG 956

Qy        431 TTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGT 467
              | |||||  |||||||||||||||||||||| |||||
Db        957 TCCACGATGCTCAACTTGGCTACTCCCTGCCGTGTGT 993


RESULT 12
AAF55259
ID   AAF55259 standard; cDNA; 3255 BP.
XX
AC   AAF55259;
XX
DT   29-MAY-2001  (first entry)
XX
```

```
DE   Nucleotide sequence of a mutated murine parkin2 polypeptide.
XX
KW   Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW   Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW   Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW   brain tumour; head trauma; stroke; vascular irregularity;
KW   metabolic irregularity; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             129. .1523
FT                   /*tag= a
FT                   /product= "truncated parkin2"
XX
PN   EP1081225-A1.
XX
PD   07-MAR-2001.
XX
PF   30-AUG-1999;   99EP-00116766.
XX
PR   30-AUG-1999;   99EP-00116766.
XX
PA   (BIOF-) BIOFRONTERA PHARM GMBH.
XX
PI   Luebbert H;
XX
DR   WPI; 2001-212797/22.
DR   P-PSDB; AAB67532.
XX
PT   New polynucleotides encoding mouse parkin2 protein, useful for producing
PT   a transgenic non-human animal as an animal model for neurodegenerative
PT   diseases.
XX
PS   Claim 3; Page 36-38; 62pp; English.
XX
CC   The present sequence encodes a murine parkin2 polypeptide. The sequence
CC   contains the mutation Thr415Asn. Mutations or deletions in the parkin2
CC   gene cause Parkinson's disease in humans. The human parkin2 gene is
CC   located in gene region 6q25.2-27. Parkin2 polypeptides and
CC   polynucleotides are useful for analysing neurodegenerative diseases. They
CC   are also useful for testing the efficacy of the treatment of a
CC   neurodegenerative disease such as Parkinson's disease, Alzheimer's
CC   disease, Huntington's disease, amyotrophic lateral sclerosis, Multi-
CC   system atrophy, Wilson's disease, Pick's disease, Prion disease, and
CC   secondary causes inducing Parkinson's syndromes like toxins, drugs, brain
CC   tumours, head trauma, stroke, vascular irregularities or metabolic
CC   irregularities, associated with a less active or non-active parkin
CC   protein
XX
SQ   Sequence 3255 BP; 853 A; 817 C; 809 G; 776 T; 0 U; 0 Other;

  Query Match             72.2%;  Score 340.2;  DB 4;  Length 3255;
  Best Local Similarity   84.0%;  Pred. No. 7.2e-107;
  Matches  384;  Conservative    0;  Mismatches   73;  Indels    0;  Gaps    0;

Qy         11 CAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGTCAAAGAGTGC 70
              ||| || | | | ||||||||||||||||    | || ||||||||||||||| || |  || |
Db        538 CAGTTAAACCCACCTACAACAGCTTTTTCATCTACTGCAAAGGCCCCTGCCACAAGGTCC 597

Qy         71 AGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTCACCTTGACCC 130
              |||| ||||| ||| | || ||||  |||||||||  || |||||  ||||||||| |||
Db        598 AGCCTGGAAAGCTCCGAGTTCAGTGTGGCACCTGCAAACAAGCAACCCTCACCTTGGCCC 657

Qy        131 AGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGTGAATGCCAAT 190
              |||| |||||||||||||| ||||| |||||||||||||||||||||||||||| ||||| |
Db        658 AGGGCCCATCTTGCTGGGACGATGTCTTAATTCCAAACCGGATGAGTGGTGAGTGCCAGT 717

Qy        191 CCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCACACCCCACCT 250
              | ||| |||||||||| || || || ||||||||||||||||||||||||||||| ||||
Db        718 CTCCAGACTGCCCTGGAACCAGAGCTGAATTTTTCTTTAAATGTGGAGCACACCCAACCT 777

Qy        251 CTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGGAACATCACTT 310
              | |||||||| || || ||||||||| |||||||| | |  || || || | |||| |||
Db        778 CAGACAAGGACACGTCGGTAGCTTTGAACCTGATCACCAGCAACAGGCGCAGCATCCCTT 837

Qy        311 GCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAACTCCCGCCACG 370
              |||| |||||||||| |||||||||| |||||||| |||||||| |||  ||| ||||
Db        838 GCATAGCGTGCACAGATGTCAGGAGCCCTGTCCTGGTCTTCCAGTGTAACCACCGTCACG 897

Qy        371 TGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGATCGGCAGTTTG 430
              |||| || || |||||||||||||| || ||||| |||||||||| ||||||||||||
Db        898 TGATCTGTTTGGACTGTTTCCACTTGTATTGTGTCACAAGACTCAACGATCGGCAGTTTG 957

Qy        431 TTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGT 467
              | |||||  |||||||||||||||||||||| |||||
Db        958 TCCACGATGCTCAACTTGGCTACTCCCTGCCGTGTGT 994


RESULT 13
AAF55244
ID   AAF55244 standard; cDNA; 3255 BP.
XX
AC   AAF55244;
XX
DT   29-MAY-2001  (first entry)
XX
DE   Nucleotide sequence of a murine parkin2 polypeptide.
XX
KW   Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW   Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW   Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW   brain tumour; head trauma; stroke; vascular irregularity;
KW   metabolic irregularity; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             129. .1523
FT                   /*tag= a
FT                   /product= "parkin2"
XX
PN   EP1081225-A1.
XX
PD   07-MAR-2001.
XX
PF   30-AUG-1999;   99EP-00116766.
XX
PR   30-AUG-1999;   99EP-00116766.
XX
PA   (BIOF-) BIOFRONTERA PHARM GMBH.
XX
PI   Luebbert H;
XX
DR   WPI; 2001-212797/22.
DR   P-PSDB; AAB67517.
XX
PT   New polynucleotides encoding mouse parkin2 protein, useful for producing
PT   a transgenic non-human animal as an animal model for neurodegenerative
PT   diseases.
XX
PS   Disclosure; Page 15-16; 62pp; English.
XX
CC   The present sequence encodes a murine parkin2 polypeptide. Mutations or
CC   deletions in the parkin2 gene cause Parkinson's disease in humans. The
CC   human parkin2 gene is located in gene region 6q25.2-27. Parkin2
CC   polypeptides and polynucleotides are useful for analysing
CC   neurodegenerative diseases. They are also useful for testing the efficacy
CC   of the treatment of a neurodegenerative disease such as Parkinson's
CC   disease, Alzheimer's disease, Huntington's disease, amyotrophic lateral
CC   sclerosis, Multi-system atrophy, Wilson's disease, Pick's disease, Prion
CC   disease, and secondary causes inducing Parkinson's syndromes like toxins,
CC   drugs, brain tumours, head trauma, stroke, vascular irregularities or
CC   metabolic irregularities, associated with a less active or non-active
CC   parkin protein
XX
```

```
SQ   Sequence 3255 BP; 852 A; 818 C; 809 G; 776 T; 0 U; 0 Other;

  Query Match             72.2%;  Score 340.2;  DB 4;  Length 3255;
  Best Local Similarity   84.0%;  Pred. No. 7.2e-107;
  Matches  384;  Conservative    0;  Mismatches   73;  Indels    0;  Gaps    0;

Qy         11 CAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGTCAAAGAGTGC 70
              ||| || | | | ||||||||||||||    | || ||||||||||||| || |  || |
Db        538 CAGTTAAACCCACCTACAACAGCTTTTTCATCTACTGCAAAGGCCCCTGCCACAAGGTCC 597

Qy         71 AGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTCACCTTGACCC 130
              |||| ||||| ||| | || |||||  |||||||||  || |||||  ||||||||| |||
Db        598 AGCCTGGAAAGCTCCGAGTTCAGTGTGGCACCTGCAAACAAGCAACCCTCACCTTGGCCC 657

Qy        131 AGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGTGAATGCCAAT 190
              |||| |||||||||||||| ||||| |||||||||||||||||||||||||||| ||||| |
Db        658 AGGGCCCATCTTGCTGGGACGATGTCTTAATTCCAAACCGGATGAGTGGTGAGTGCCAGT 717

Qy        191 CCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCACACCCCACCT 250
              | ||| |||||||||| || || || ||||||||||||||||||||||||||||| ||||
Db        718 CTCCAGACTGCCCTGGAACCAGAGCTGAATTTTTCTTTAAATGTGGAGCACACCCAACCT 777

Qy        251 CTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGGAACATCACTT 310
              | |||||||| || || ||||||||| |||||||| | |  || || || | |||| |||
Db        778 CAGACAAGGACACGTCGGTAGCTTTGAACCTGATCACCAGCAACAGGCGCAGCATCCCTT 837

Qy        311 GCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAACTCCCGCCACG 370
              ||||  |||||||||| |||||||||| |||||||| |||||||| |||  ||| ||||
Db        838 GCATAGCGTGCACAGATGTCAGGAGCCCTGTCCTGGTCTTCCAGTGTAACCACCGTCACG 897

Qy        371 TGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGATCGGCAGTTTG 430
              |||| || || |||||||||||||| || ||||| ||||||||||| ||||||||||||
Db        898 TGATCTGTTTGGACTGTTTCCACTTGTATTGTGTCACAAGACTCAACGATCGGCAGTTTG 957

Qy        431 TTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGT 467
              | |||||  |||||||||||||||||||||| |||||
Db        958 TCCACGATGCTCAACTTGGCTACTCCCTGCCGTGTGT 994


RESULT 14
AAF55260
ID   AAF55260 standard; cDNA; 3255 BP.
XX
AC   AAF55260;
XX
DT   29-MAY-2001  (first entry)
XX
DE   Nucleotide sequence of a mutated murine parkin2 polypeptide.
XX
KW   Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW   Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW   Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW   brain tumour; head trauma; stroke; vascular irregularity;
KW   metabolic irregularity; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             129. .1484
FT                   /*tag=  a
FT                   /product= "truncated parkin2"
XX
PN   EP1081225-A1.
XX
PD   07-MAR-2001.
XX
PF   30-AUG-1999;   99EP-00116766.
XX
PR   30-AUG-1999;   99EP-00116766.
XX
PA   (BIOF-) BIOFRONTERA PHARM GMBH.
XX
PI   Luebbert H;
XX
DR   WPI; 2001-212797/22.
DR   P-PSDB; AAB67533.
XX
PT   New polynucleotides encoding mouse parkin2 protein, useful for producing
PT   a transgenic non-human animal as an animal model for neurodegenerative
PT   diseases.
XX
PS   Claim 3; Page 38-39; 62pp; English.
XX
CC   The present sequence encodes a murine parkin2 polypeptide. The sequence
CC   contains the mutation Trp453Stop. Mutations or deletions in the parkin2
CC   gene cause Parkinson's disease in humans. The human parkin2 gene is
CC   located in gene region 6q25.2-27. Parkin2 polypeptides and
CC   polynucleotides are useful for analysing neurodegenerative diseases. They
CC   are also useful for testing the efficacy of the treatment of a
CC   neurodegenerative disease such as Parkinson's disease, Alzheimer's
CC   disease, Huntington's disease, amyotrophic lateral sclerosis, Multi-
CC   system atrophy, Wilson's disease, Pick's disease, Prion disease, and
CC   secondary causes inducing Parkinson's syndromes like toxins, drugs, brain
CC   tumours, head trauma, stroke, vascular irregularities or metabolic
CC   irregularities, associated with a less active or non-active parkin
CC   protein
XX
SQ   Sequence 3255 BP; 853 A; 818 C; 808 G; 776 T; 0 U; 0 Other;

  Query Match             72.2%;  Score 340.2;  DB 4;  Length 3255;
  Best Local Similarity   84.0%;  Pred. No. 7.2e-107;
  Matches  384;  Conservative    0;  Mismatches   73;  Indels    0;  Gaps    0;

Qy         11 CAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGTCAAAGAGTGC 70
              ||| || | | | ||||||||||||||    | || ||||||||||||| || |  || |
Db        538 CAGTTAAACCCACCTACAACAGCTTTTTCATCTACTGCAAAGGCCCCTGCCACAAGGTCC 597

Qy         71 AGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTCACCTTGACCC 130
              |||| ||||| ||| | || |||||  |||||||||  || |||||  ||||||||| |||
Db        598 AGCCTGGAAAGCTCCGAGTTCAGTGTGGCACCTGCAAACAAGCAACCCTCACCTTGGCCC 657

Qy        131 AGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGTGAATGCCAAT 190
              |||| |||||||||||||| ||||| |||||||||||||||||||||||||||| ||||| |
Db        658 AGGGCCCATCTTGCTGGGACGATGTCTTAATTCCAAACCGGATGAGTGGTGAGTGCCAGT 717

Qy        191 CCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCACACCCCACCT 250
              | ||| |||||||||| || || || ||||||||||||||||||||||||||||| ||||
Db        718 CTCCAGACTGCCCTGGAACCAGAGCTGAATTTTTCTTTAAATGTGGAGCACACCCAACCT 777

Qy        251 CTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGGAACATCACTT 310
              | |||||||| || || ||||||||| |||||||| | |  || || || | |||| |||
Db        778 CAGACAAGGACACGTCGGTAGCTTTGAACCTGATCACCAGCAACAGGCGCAGCATCCCTT 837

Qy        311 GCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAACTCCCGCCACG 370
              ||||  |||||||||| |||||||||| |||||||| |||||||| |||  ||| ||||
Db        838 GCATAGCGTGCACAGATGTCAGGAGCCCTGTCCTGGTCTTCCAGTGTAACCACCGTCACG 897

Qy        371 TGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGATCGGCAGTTTG 430
              |||| || || |||||||||||||| || ||||| ||||||||||| ||||||||||||
Db        898 TGATCTGTTTGGACTGTTTCCACTTGTATTGTGTCACAAGACTCAACGATCGGCAGTTTG 957

Qy        431 TTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGT 467
              | |||||  |||||||||||||||||||||| |||||
Db        958 TCCACGATGCTCAACTTGGCTACTCCCTGCCGTGTGT 994


RESULT 15
AAF55258
ID   AAF55258 standard; cDNA; 3255 BP.
XX
AC   AAF55258;
XX
DT   29-MAY-2001  (first entry)
XX
DE   Nucleotide sequence of a mutated murine parkin2 polypeptide.
XX
```

```
KW   Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW   Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW   Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW   brain tumour; head trauma; stroke; vascular irregularity;
KW   metabolic irregularity; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             129. .1523
FT                   /*tag= a
FT                   /product= "truncated parkin2"
XX
PN   EP1081225-A1.
XX
PD   07-MAR-2001.
XX
PF   30-AUG-1999;   99EP-00116766.
XX
PR   30-AUG-1999;   99EP-00116766.
XX
PA   (BIOF-) BIOFRONTERA PHARM GMBH.
XX
PI   Luebbert H;
XX
DR   WPI; 2001-212797/22.
DR   P-PSDB; AAB67531.
XX
PT   New polynucleotides encoding mouse parkin2 protein, useful for producing
PT   a transgenic non-human animal as an animal model for neurcdegenerative
PT   diseases.
XX
PS   Claim 3; Page 35-36; 62pp; English.
XX
CC   The present sequence encodes a murine parkin2 polypeptide. The sequence
CC   contains the mutation Lys161Asn. Mutations or deletions in the parkin2
CC   gene cause Parkinson's disease in humans. The human parkin2 gene is
CC   located in gene region 6q25.2-27. Parkin2 polypeptides and
CC   polynucleotides are useful for analysing neurodegenerative diseases. They
CC   are also useful for testing the efficacy of the treatment of a
CC   neurodegenerative disease such as Parkinson's disease, Alzheimer's
CC   disease, Huntington's disease, amyotrophic lateral sclerosis, Multi-
CC   system atrophy, Wilson's disease, Pick's disease, Prion disease, and
CC   secondary causes inducing Parkinson's syndromes like toxins, drugs, brain
CC   tumours, head trauma, stroke, vascular irregularities or metabolic
CC   irregularities, associated with a less active or non-active parkin
CC   protein
XX
SQ   Sequence 3255 BP; 852 A; 818 C; 808 G; 777 T; 0 U; 0 Other;

  Query Match             72.2%;  Score 340.2;  DB 4;  Length 3255;
  Best Local Similarity   84.0%;  Pred. No. 7.2e-107;
  Matches  384;  Conservative    0;  Mismatches   73;  Indels    0;  Gaps    0;

Qy         11 CAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGTCAAAGAGTGC 70
              ||| || | | | ||||||||||||||    | || ||||||||||||| || |  || |
Db        538 CAGTTAAACCCACCTACAACAGCTTTTTCATCTACTGCAAAGGCCCCTGCCACAAGGTCC 597

Qy         71 AGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTCACCTTGACCC 130
              |||| ||||| ||| | || |||||  |||||||||  || |||||  ||||||||| |||
Db        598 AGCCTGGAAATCTCCGAGTTCAGTGTGGCACCTGCAAACAAGCAACCCTCACCTTGGCCC 657

Qy        131 AGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGTGAATGCCAAT 190
              |||| |||||||||||||| ||||| |||||||||||||||||||||||||||| ||||| |
Db        658 AGGGCCCATCTTGCTGGGACGATGTCTTAATTCCAAACCGGATGAGTGGTGAGTGCCAGT 717

Qy        191 CCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCACACCCCACCT 250
              | ||| |||||||||| || || || ||||||||||||||||||||||||||||| ||||
Db        718 CTCCAGACTGCCCTGGAACCAGAGCTGAATTTTTCTTTAAATGTGGAGCACACCCAACCT 777

Qy        251 CTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGGAACATCACTT 310
              | ||||||||| || || ||||||||| |||||||| | |  || || || | |||| |||
Db        778 CAGACAAGGACACGTCGGTAGCTTTGAACCTGATCACCAGCAACAGGCGCAGCATCCCTT 837

Qy        311 GCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAACTCCCGCCACG 370
              ||||  |||||||||| |||||||||| |||||||| |||||||| |||  ||| ||||
Db        838 GCATAGCGTGCACAGATGTCAGGAGCCCTGTCCTGGTCTTCCAGTGTAACCACCGTCACG 897

Qy        371 TGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGATCGGCAGTTTG 430
              |||| || || ||||||||||||| || ||||| |||||||||| ||||||||||||
Db        898 TGATCTGTTTGGACTGTTTCCACTTGTATTGTGTCACAAGACTCAACGATCGGCAGTTTG 957

Qy        431 TTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGT 467
              | |||||  |||||||||||||||||||||| |||||
Db        958 TCCACGATGCTCAACTTGGCTACTCCCTGCCGTGTGT 994

Search completed: February 25, 2005, 19:01:27
Job time : 268.776 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 25, 2005, 17:58:29 ; Search time 81.1144 Seconds
(without alignments)
9501.226 Million cell updates/sec

Title: US-10-622-817-10
Perfect score: 471
Sequence: 1 ggaagtccagcaggtagatc.....actccctgccttgtgtgtag 471

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	466.4	99.0	2960	4	US-09-601-844B-1	Sequence 1, Appli
	2	466.4	99.0	2960	4	US-09-949-016-645	Sequence 645, App
	3	288.4	61.2	2876	4	US-09-601-844B-3	Sequence 3, Appli
	4	288.4	61.2	2876	4	US-09-949-016-982	Sequence 982, App
	5	251.8	53.5	2513	4	US-09-949-016-983	Sequence 983, App
	6	37.8	8.0	399	4	US-09-621-976-8976	Sequence 8976, Ap
	7	34.8	7.4	40	4	US-09-601-844B-65	Sequence 65, Appl
	8	34.8	7.4	361	3	US-09-643-597-293	Sequence 293, App
	9	34.8	7.4	361	4	US-09-480-884A-293	Sequence 293, App
	10	34.8	7.4	361	4	US-09-542-615A-293	Sequence 293, App
	11	34.8	7.4	361	4	US-09-606-421B-293	Sequence 293, App
	12	34.8	7.4	361	4	US-09-630-940B-293	Sequence 293, App
c	13	34.8	7.4	405	4	US-09-513-999C-11324	Sequence 11324, A
	14	34	7.2	40	4	US-09-601-844B-64	Sequence 64, Appl
c	15	32.8	7.0	178883	4	US-09-949-016-12733	Sequence 12733, A
c	16	32.8	7.0	178884	4	US-09-949-016-13039	Sequence 13039, A
	17	32.6	6.9	35262	4	US-09-949-016-12797	Sequence 12797, A
	18	32.6	6.9	35263	4	US-09-949-016-16399	Sequence 16399, A
c	19	32.2	6.8	152132	4	US-09-949-016-13845	Sequence 13845, A
c	20	32.2	6.8	152145	4	US-09-949-016-12371	Sequence 12371, A
	21	32	6.8	38	4	US-09-601-844B-66	Sequence 66, Appl
c	22	32	6.8	347	4	US-09-621-976-11665	Sequence 11665, A
c	23	32	6.8	819	4	US-09-489-039A-3239	Sequence 3239, Ap
	24	31.8	6.8	37	4	US-09-601-844B-68	Sequence 68, Appl
	25	31.8	6.8	342	4	US-09-134-000C-2019	Sequence 2019, Ap
	26	31.8	6.8	1160	3	US-09-737-698B-12	Sequence 12, Appl
	27	31.8	6.8	1160	4	US-09-737-626A-12	Sequence 12, Appl
	28	31.8	6.8	1695	3	US-09-737-698B-28	Sequence 28, Appl
	29	31.8	6.8	1695	4	US-09-737-626A-28	Sequence 28, Appl
c	30	31.8	6.8	1830121	4	US-09-557-884-1	Sequence 1, Appli
c	31	31.8	6.8	1830121	4	US-09-643-990A-1	Sequence 1, Appli
	32	31.6	6.7	601	4	US-09-949-016-36816	Sequence 36816, A
	33	31.6	6.7	601	4	US-09-949-016-58953	Sequence 58953, A
	34	31.6	6.7	601	4	US-09-949-016-184559	Sequence 184559,
c	35	31.6	6.7	119801	4	US-09-949-016-13453	Sequence 13453, A
c	36	31.6	6.7	190078	4	US-09-949-016-12707	Sequence 12707, A
c	37	31.6	6.7	190078	4	US-09-949-016-17026	Sequence 17026, A
	38	31.6	6.7	312470	4	US-09-949-016-14043	Sequence 14043, A
	39	31.6	6.7	336024	4	US-09-949-016-12373	Sequence 12373, A
c	40	31.4	6.7	832	4	US-09-621-976-2813	Sequence 2813, Ap
	41	31.2	6.6	27727	4	US-09-949-016-15737	Sequence 15737, A
	42	31.2	6.6	29067	4	US-09-949-016-12535	Sequence 12535, A
	43	31.2	6.6	29093	4	US-09-949-016-17476	Sequence 17476, A
c	44	31	6.6	74940	4	US-09-949-016-17512	Sequence 17512, A
c	45	30.8	6.5	7194	4	US-09-601-326-76	Sequence 76, Appl

ALIGNMENTS

RESULT 1
US-09-601-844B-1
; Sequence 1, Application US/09601844B
; Patent No. 6716621
; GENERAL INFORMATION:
; APPLICANT: Shimizu, No. 6716621uyoshi
; APPLICANT: Mizuno, Yoshikuni
; TITLE OF INVENTION: Isolated DNA or Gene Responsible for Parkinson's Disease
; FILE REFERENCE: 0652.2110000
; CURRENT APPLICATION NUMBER: US/09/601,844B
; CURRENT FILING DATE: 2000-08-09
; PRIOR APPLICATION NUMBER: PCT/JP99/00545
; PRIOR FILING DATE: 1999-02-09
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2960
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (102)..(1496)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (102)..(108)
; OTHER INFORMATION: Exon 1
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (109)..(272)
; OTHER INFORMATION: Exon 2
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (273)..(513)
; OTHER INFORMATION: Exon 3
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (514)..(635)
; OTHER INFORMATION: Exon 4
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (636)..(719)
; OTHER INFORMATION: Exon 5
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (720)..(835)
; OTHER INFORMATION: Exon 6
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (836)..(972)

```
;   OTHER INFORMATION: Exon 7
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (973)..(1034)
;   OTHER INFORMATION: Exon 8
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1035)..(1184)
;   OTHER INFORMATION: Exon 9
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1185)..(1268)
;   OTHER INFORMATION: Exon 10
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1269)..(1386)
;   OTHER INFORMATION: Exon 11
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1387)..(2960)
;   OTHER INFORMATION: Exon 12
US-09-601-844B-1

  Query Match             99.0%;  Score 466.4;  DB 4;  Length 2960;
  Best Local Similarity   99.8%;  Pred. No. 1.9e-154;
  Matches  467;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 563

Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 623

Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 683

Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 743

Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        744 CACCCCACCTCTGACAAGGAAACACCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 803

Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 863

Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 923

Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 971


RESULT 2
US-09-949-016-645
; Sequence 645, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 645
;   LENGTH: 2960
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-645

  Query Match             99.0%;  Score 466.4;  DB 4;  Length 2960;
  Best Local Similarity   99.8%;  Pred. No. 1.9e-154;
  Matches  467;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 563

Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 623

Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 683

Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 743

Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        744 CACCCCACCTCTGACAAGGAAACACCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 803

Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 863

Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 923

Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 971


RESULT 3
US-09-601-844B-3
; Sequence 3, Application US/09601844B
; Patent No. 6716621
; GENERAL INFORMATION:
;  APPLICANT: Shimizu, No. 6716621uyoshi
;  APPLICANT:  Mizuno, Yoshikuni
;  TITLE OF INVENTION: Isolated DNA or Gene Responsible for Parkinson's Disease
;  FILE REFERENCE: 0652.2110000
;  CURRENT APPLICATION NUMBER: US/09/601,844B
;  CURRENT FILING DATE:  2000-08-09
;  PRIOR APPLICATION NUMBER: PCT/JP99/00545
;  PRIOR FILING DATE: 1999-02-09
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 2876
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (102)..(1412)
;   OTHER INFORMATION:
;   FEATURE:
```

```
;   NAME/KEY: misc_feature
;   LOCATION: (102)..(108)
;   OTHER INFORMATION: Exon 1
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (109)..(272)
;   OTHER INFORMATION: Exon 2
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (273)..(513)
;   OTHER INFORMATION: Exon 3
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (514)..(635)
;   OTHER INFORMATION: Exon 4
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (636)..(751)
;   OTHER INFORMATION: Exon 6
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (752)..(888)
;   OTHER INFORMATION: Exon 7
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (889)..(950)
;   OTHER INFORMATION: Exon 8
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (951)..(1100)
;   OTHER INFORMATION: Exon 9
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1101)..(1184)
;   OTHER INFORMATION: Exon 10
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1185)..(1302)
;   OTHER INFORMATION: Exon 11
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1303)..(2876)
;   OTHER INFORMATION: Exon 11
US-09-601-844B-3

  Query Match             61.2%;  Score 288.4;  DB 4;  Length 2876;
  Best Local Similarity   81.8%;  Pred. No. 2.7e-91;
  Matches  383;  Conservative    0;  Mismatches    1;  Indels   84;  Gaps    1;

Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 563

Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 623

Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
              ||||||||||||
Db        624 ACCTTGACCCAG---------------------------------------------- 635

Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
                                                  ||||||||||||||||||||||||
Db        636 ----------------------------------GAATTTTTCTTTAAATGTGGAGCA 659

Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        660 CACCCCACCTCTGACAAGGAAACACCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 719

Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 779

Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 839

Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        840 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 887


RESULT 4
US-09-949-016-982
; Sequence 982, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 982
;   LENGTH: 2876
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-982

  Query Match             61.2%;  Score 288.4;  DB 4;  Length 2876;
  Best Local Similarity   81.8%;  Pred. No. 2.7e-91;
  Matches  383;  Conservative    0;  Mismatches    1;  Indels   84;  Gaps    1;

Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 563

Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 623

Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
              ||||||||||||
Db        624 ACCTTGACCCAG---------------------------------------------- 635

Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
                                                  ||||||||||||||||||||||||
Db        636 ----------------------------------GAATTTTTCTTTAAATGTGGAGCA 659

Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        660 CACCCCACCTCTGACAAGGAAACACCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 719

Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 779

Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 839

Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        840 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 887


RESULT 5
US-09-949-016-983
```

```
; Sequence 983, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 983
;   LENGTH: 2513
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-983

  Query Match             53.5%;  Score 251.8;  DB 4;  Length 2513;
  Best Local Similarity   97.3%;  Pred. No. 2.4e-78;
  Matches  256;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy        206 GGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCACACCCCACCTCTGACAAGGAAACAT 265
              |||||     ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GGACTGTGCAGGAATTTTTCTTTAAATGTGGAGCACACCCCACCTCTGACAAGGAAACAC 321

Qy        266 CAGTAGCTTTGCACCTGATCGCAACAAATAGTCGGAACATCACTTGCATTACGTGCACAG 325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 CAGTAGCTTTGCACCTGATCGCAACAAATAGTCGGAACATCACTTGCATTACGTGCACAG 381

Qy        326 ACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAACTCCCGCCACGTGATTTGCTTAGACT 385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        382 ACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAACTCCCGCCACGTGATTTGCTTAGACT 441

Qy        386 GTTTCCACTTATACTGTGTGACAAGACTCAATGATCGGCAGTTTGTTCACGACCCTCAAC 445
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        442 GTTTCCACTTATACTGTGTGACAAGACTCAATGATCGGCAGTTTGTTCACGACCCTCAAC 501

Qy        446 TTGGCTACTCCCTGCCTTGTGTG 468
              |||||||||||||||||||||||
Db        502 TTGGCTACTCCCTGCCTTGTGTG 524


RESULT 6
US-09-621-976-8976
; Sequence 8976, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 8976
;   LENGTH: 399
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-621-976-8976

  Query Match              8.0%;  Score 37.8;  DB 4;  Length 399;
  Best Local Similarity   12.0%;  Pred. No. 0.0069;
  Matches   36;  Conservative  134;  Mismatches  131;  Indels    0;  Gaps    0;

Qy        166 AACCGGATGAGTGGTGAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTC 225
              :|: |  |: ::::::|  : :       : :::  ::   : ::  ::::|::|:  :
Db         20 RAYWGMYTRGSKKSWRAKMSWMKKRRRRRRAWMWWKSMCWMKKSKSWWRSWGMMTKRMKG 79

Qy        226 TTTAAATGTGGAGCACACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATC 285
              | : |: ::   :      : : ::||   :|| ::  ||:| : : ::::  :
Db         80 RGAASWAGYMSWMTYMTRRWRYYRYRKKACTKWRAAGMWGKAGWAWMAYAKMWYMAWRRT 139

Qy        286 GCAACAAATAGTCGGAACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTG 345
              :   |::: : : ::: |::|  : ::  :  :|::     | :: :||:::   ::
Db        140 AMKYWAMMKSKRSMRRRRAWYAWMYYMMARRTMWGMRASCYRGAYMASAGMYYWMMYYMM 199

Qy        346 GTTTTCCAGTGCAACTCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTG 405
              : :: ::||:::::  :     : : :: :     :: ::::|   :  :  ::||::: :
Db        200 RRKWMYSAGWSMMRKWTRRCASYSCWSSYCMWGAKMMMYWKTSRWSYWYSSYRCTKYRRS 259

Qy        406 ACAAGACTCAATGATCGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGT 465
              | :::: :: |   :  ||: :::::     ||: |  :::::: |: ::::  : ::|
Db        260 CCCWSMSCYWKTYYRSWYCASCYYSYKTKRASCCMMCCCMKWRKMMAMMYMMKMTYCKKT 319

Qy        466 G 466
              :
Db        320 S 320


RESULT 7
US-09-601-844B-65
; Sequence 65, Application US/09601844B
; Patent No. 6716621
; GENERAL INFORMATION:
;  APPLICANT: Shimizu, No. 6716621uyoshi
;  APPLICANT:  Mizuno, Yoshikuni
;  TITLE OF INVENTION: Isolated DNA or Gene Responsible for Parkinson's Disease
;  FILE REFERENCE: 0652.2110000
;  CURRENT APPLICATION NUMBER: US/09/601,844B
;  CURRENT FILING DATE:  2000-08-09
;  PRIOR APPLICATION NUMBER: PCT/JP99/00545
;  PRIOR FILING DATE: 1999-02-09
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 65
;   LENGTH: 40
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-601-844B-65

  Query Match              7.4%;  Score 34.8;  DB 4;  Length 40;
  Best Local Similarity   94.7%;  Pred. No. 0.02;
  Matches   36;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        128 CCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCA 165
              || | ||||||||||||||||||||||||||||||||||
Db          3 CCAAAGGTCCATCTTGCTGGGATGATGTTTTAATTCCA 40


RESULT 8
US-09-643-597-293
; Sequence 293, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  McNeill, Patricia D.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.455C11
```

```
;  CURRENT APPLICATION NUMBER: US/09/643,597
;  CURRENT FILING DATE:  2000-08-21
;  NUMBER OF SEQ ID NOS:  369
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 293
;   LENGTH: 361
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(361)
;   OTHER INFORMATION: n = A,T,C or G
US-09-643-597-293

  Query Match             7.4%;  Score 34.8;  DB 3;  Length 361;
  Best Local Similarity   59.4%;  Pred. No. 0.076;
  Matches   57;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy        203 CTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCACACCCCACCTCTGACAAGGAAA 262
              || | ||  ||    ||||||| |||  ||  ||  |||  || ||| |  |  || |||
Db        134 CTTGAACACGTCTTCAATTTTCCTTCCAAAATGCTGCATGCCACACTTGAGGTAACGAAG 193

Qy        263 CATCAGTAGCTTTGCACCTGATCGCAACAAATAGTC 298
              ||  ||||  |||  || |||  || |  || | ||
Db        194 CANAAGTATTTTTAAACATGACAGCTAANAACATTC 229


RESULT 9
US-09-480-884A-293
; Sequence 293, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Hosken, Nancy A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C6
;  CURRENT APPLICATION NUMBER: US/09/480,884A
;  CURRENT FILING DATE:  2001-08-27
;  NUMBER OF SEQ ID NOS: 330
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 293
;   LENGTH: 361
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(361)
;   OTHER INFORMATION: n = A,T,C or G
US-09-480-884A-293

  Query Match             7.4%;  Score 34.8;  DB 4;  Length 361;
  Best Local Similarity   59.4%;  Pred. No. 0.076;
  Matches   57;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy        203 CTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCACACCCCACCTCTGACAAGGAAA 262
              || | ||  ||    ||||||| |||  ||  ||  |||  || ||| |  |  || |||
Db        134 CTTGAACACGTCTTCAATTTTCCTTCCAAAATGCTGCATGCCACACTTGAGGTAACGAAG 193

Qy        263 CATCAGTAGCTTTGCACCTGATCGCAACAAATAGTC 298
              ||  ||||  |||  || |||  || |  || | ||
Db        194 CANAAGTATTTTTAAACATGACAGCTAANAACATTC 229


RESULT 10
US-09-542-615A-293
; Sequence 293, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy A.
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.455C8
;  CURRENT APPLICATION NUMBER: US/09/542,615A
;  CURRENT FILING DATE:  2000-04-14
;  NUMBER OF SEQ ID NOS:  350
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 293
;   LENGTH: 361
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(361)
;   OTHER INFORMATION: n = A,T,C or G
US-09-542-615A-293

  Query Match             7.4%;  Score 34.8;  DB 4;  Length 361;
  Best Local Similarity   59.4%;  Pred. No. 0.076;
  Matches   57;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy        203 CTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCACACCCCACCTCTGACAAGGAAA 262
              || | ||  ||    ||||||| |||  ||  ||  |||  || ||| |  |  || |||
Db        134 CTTGAACACGTCTTCAATTTTCCTTCCAAAATGCTGCATGCCACACTTGAGGTAACGAAG 193

Qy        263 CATCAGTAGCTTTGCACCTGATCGCAACAAATAGTC 298
              ||  ||||  |||  || |||  || |  || | ||
Db        194 CANAAGTATTTTTAAACATGACAGCTAANAACATTC 229


RESULT 11
US-09-606-421B-293
; Sequence 293, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C9
;  CURRENT APPLICATION NUMBER: US/09/606,421B
;  CURRENT FILING DATE:  2000-06-28
;  NUMBER OF SEQ ID NOS: 358
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 293
;   LENGTH: 361
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(361)
;   OTHER INFORMATION: n = A,T,C or G
US-09-606-421B-293

  Query Match             7.4%;  Score 34.8;  DB 4;  Length 361;
  Best Local Similarity   59.4%;  Pred. No. 0.076;
  Matches   57;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy        203 CTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCACACCCCACCTCTGACAAGGAAA 262
              || | ||  ||    ||||||| |||  ||  ||  |||  || ||| |  |  || |||
Db        134 CTTGAACACGTCTTCAATTTTCCTTCCAAAATGCTGCATGCCACACTTGAGGTAACGAAG 193
```

```
Qy        263 CATCAGTAGCTTTGCACCTGATCGCAACAAATAGTC 298
              ||  ||||  |||  ||  |||  ||  |  || | ||
Db        194 CANAAGTATTTTTAAACATGACAGCTAANAACATTC 229


RESULT 12
US-09-630-940B-293
; Sequence 293, Application US/09630940B
; Patent No. 6737514
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  McNeill, Patricia D.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C10
;  CURRENT APPLICATION NUMBER: US/09/630,940B
;  CURRENT FILING DATE:  2000-08-02
;  NUMBER OF SEQ ID NOS: 367
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 293
;   LENGTH: 361
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(361)
;   OTHER INFORMATION: n = A,T,C or G
US-09-630-940B-293

  Query Match             7.4%;  Score 34.8;  DB 4;  Length 361;
  Best Local Similarity   59.4%;  Pred. No. 0.076;
  Matches   57;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy        203 CTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCACACCCCACCTCTGACAAGGAAA 262
              || | ||  ||    ||||||| |||  ||  ||  |||  || ||| |  |  || |||
Db        134 CTTGAACACGTCTTCAATTTTCCTTCCAAAATGCTGCATGCCACACTTGAGGTAACGAAG 193

Qy        263 CATCAGTAGCTTTGCACCTGATCGCAACAAATAGTC 298
              ||  ||||  |||  ||  |||  ||  |  || | ||
Db        194 CANAAGTATTTTTAAACATGACAGCTAANAACATTC 229


RESULT 13
US-09-513-999C-11324/c
; Sequence 11324, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 11324
;   LENGTH: 405
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 326
;   OTHER INFORMATION: w=a or t
US-09-513-999C-11324

  Query Match             7.4%;  Score 34.8;  DB 4;  Length 405;
  Best Local Similarity   59.4%;  Pred. No. 0.081;
  Matches   57;  Conservative    1;  Mismatches   38;  Indels    0;  Gaps    0;

Qy        203 CTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCACACCCCACCTCTGACAAGGAAA 262
              || | ||: ||    ||||||| |||  ||  ||  |||  || ||| |  |  || |||
Db        333 CTTGAACWCGTCTTCAATTTTCCTTCCAAAATGCTGCATGCCACACTTGAGGTAACGAAG 274

Qy        263 CATCAGTAGCTTTGCACCTGATCGCAACAAATAGTC 298
              ||  ||||  |||  ||  |||  ||  |  || | ||
Db        273 CAGAAGTATTTTTAAACATGACAGCTAAGAACATTC 238


RESULT 14
US-09-601-844B-64
; Sequence 64, Application US/09601844B
; Patent No. 6716621
; GENERAL INFORMATION:
;  APPLICANT: Shimizu, No. 6716621uyoshi
;  APPLICANT:  Mizuno, Yoshikuni
;  TITLE OF INVENTION: Isolated DNA or Gene Responsible for Parkinson's Disease
;  FILE REFERENCE: 0652.2110000
;  CURRENT APPLICATION NUMBER: US/09/601,844B
;  CURRENT FILING DATE:  2000-08-09
;  PRIOR APPLICATION NUMBER: PCT/JP99/00545
;  PRIOR FILING DATE: 1999-02-09
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64
;   LENGTH: 40
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-601-844B-64

  Query Match             7.2%;  Score 34;  DB 4;  Length 40;
  Best Local Similarity   100.0%;  Pred. No. 0.038;
  Matches   34;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        131 AGGGTCCATCTTGCTGGGATGATGTTTTAATTCC 164
              ||||||||||||||||||||||||||||||||||
Db          7 AGGGTCCATCTTGCTGGGATGATGTTTTAATTCC 40


RESULT 15
US-09-949-016-12733/c
; Sequence 12733, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12733
;   LENGTH: 178883
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12733
```

Query Match 7.0%; Score 32.8; DB 4; Length 178883;
Best Local Similarity 54.0%; Pred. No. 17;
Matches 67; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

```
Qy         33 CTTTTATGTGTATTGCAAAGGCCCCTGTCAAAGAGTGCAGCCGGGAAAACTCAGGGTACA 92
              ||| |  ||||  ||||  |     | |  ||  |||   |   |  | |   |  ||||
Db     109301 CTTGTCCGTGTGATGCATTGCTGGGTTTGCAAATGTGACTCACAGCCAGCAGGGCTTACA 109242

Qy         93 GTGCAGCACCTGCAGGCAGGCAACGCTCACCTTGACCCAGGGTCCATCTTGCTGGGATGA 152
                ||   ||||||| ||  |||  ||| |  ||  |||||  ||  |||| | | ||
Db     109241 AACTAGTGTCTGCAGGGAGAGAACAGTCATCCAGAATCAGGGCACAGTTTGCGGTGGTGC 109182

Qy        153 TGTT 156
              || |
Db     109181 TGCT 109178
```

Search completed: February 26, 2005, 05:20:30
Job time : 88.1144 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 26, 2005, 02:13:05 ; Search time 283.738 Seconds
(without alignments)
9836.161 Million cell updates/sec

Title: US-10-622-817-10
Perfect score: 471
Sequence: 1 ggaagtccagcaggtagatc.....actccctgccttgtgtgtag 471

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5394803 seqs, 2962729879 residues

Total number of hits satisfying chosen parameters: 10789606

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	471	100.0	471	9	US-09-785-548-3	Sequence 3, Appli
2	471	100.0	471	18	US-10-622-817-10	Sequence 10, Appl
3	466.4	99.0	2955	19	US-10-839-688-12	Sequence 12, Appl
4	466.4	99.0	2960	18	US-10-473-226-1	Sequence 1, Appli
5	466.4	99.0	2960	18	US-10-473-226-3	Sequence 3, Appli
6	466.4	99.0	2960	18	US-10-622-817-7	Sequence 7, Appli
7	466.4	99.0	2960	18	US-10-776-604-1	Sequence 1, Appli
8	466.4	99.0	2960	19	US-10-839-688-11	Sequence 11, Appl
9	347	73.7	1470	18	US-10-622-817-13	Sequence 13, Appl
10	288.4	61.2	2876	18	US-10-776-604-3	Sequence 3, Appli
11	138.6	29.4	650	19	US-10-839-688-6	Sequence 6, Appli
12	89.8	19.1	660	19	US-10-839-688-4	Sequence 4, Appli
13	50.2	10.7	65	10	US-09-908-975-26043	Sequence 26043, A
14	34.8	7.4	40	18	US-10-776-604-65	Sequence 65, Appl
15	34.8	7.4	361	9	US-09-735-705-293	Sequence 293, App
16	34.8	7.4	361	9	US-09-850-716A-293	Sequence 293, App
17	34.8	7.4	361	9	US-09-897-778-293	Sequence 293, App
18	34.8	7.4	361	14	US-10-007-700-293	Sequence 293, App
19	34.8	7.4	361	15	US-10-117-982-293	Sequence 293, App
20	34.8	7.4	361	17	US-10-313-986-293	Sequence 293, App
21	34.8	7.4	361	18	US-10-775-972-293	Sequence 293, App
c 22	34.6	7.3	339	19	US-10-696-639-819	Sequence 819, App
23	34	7.2	40	18	US-10-776-604-64	Sequence 64, Appl
c 24	33.6	7.1	203	17	US-10-242-535A-10255	Sequence 10255, A
c 25	33.6	7.1	203	17	US-10-085-783A-10255	Sequence 10255, A
c 26	33.6	7.1	389	17	US-10-242-535A-12486	Sequence 12486, A
c 27	33.6	7.1	389	17	US-10-085-783A-12486	Sequence 12486, A
c 28	33.6	7.1	396	11	US-09-969-034-578	Sequence 578, App
c 29	33.6	7.1	447	10	US-09-918-995-14121	Sequence 14121, A
c 30	33.6	7.1	483	10	US-09-918-995-13976	Sequence 13976, A
c 31	33.6	7.1	674	16	US-10-076-747-43	Sequence 43, Appl
c 32	33.6	7.1	674	17	US-10-001-885-78	Sequence 78, Appl
c 33	33.6	7.1	676	11	US-09-989-890-61	Sequence 61, Appl
c 34	33.6	7.1	1344	17	US-10-001-885-92	Sequence 92, Appl
c 35	33.6	7.1	1375	17	US-10-001-885-79	Sequence 79, Appl
c 36	33.6	7.1	1756	13	US-10-001-887-81	Sequence 81, Appl
c 37	33.6	7.1	2731748	18	US-10-297-465A-1	Sequence 1, Appli
38	33	7.0	149382	19	US-10-741-600-17661	Sequence 17661, A
39	32.8	7.0	4251	16	US-10-240-965-106	Sequence 106, App
40	32.8	7.0	41694	18	US-10-719-993-6806	Sequence 6806, Ap
c 41	32.8	7.0	194883	13	US-10-087-192-826	Sequence 826, App
c 42	32.2	6.8	657	18	US-10-767-701-2873	Sequence 2873, Ap
43	32.2	6.8	1119	17	US-10-242-355-195	Sequence 195, App
c 44	32.2	6.8	1738	18	US-10-739-930-475	Sequence 475, App
45	32.2	6.8	3148	18	US-10-437-963-80145	Sequence 80145, A

ALIGNMENTS

```
RESULT 1
US-09-785-548-3
; Sequence 3, Application US/09785548
; Patent No. US20020155577A1
; GENERAL INFORMATION:
;  APPLICANT: AVENTIS PHARMACEUTICALS, INC.
;  TITLE OF INVENTION: COMPOSITIONS THAT CAN BE USED FOR REGULATING THE ACTIVITY OF PARK
;  FILE REFERENCE: ST00005
;  CURRENT APPLICATION NUMBER: US/09/785,548
;  CURRENT FILING DATE:  2001-02-20
;  NUMBER OF SEQ ID NOS: 46
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 471
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(471)
US-09-785-548-3

  Query Match           100.0%;  Score 471;  DB 9;  Length 471;
  Best Local Similarity   100.0%;  Pred. No. 9.5e-157;
  Matches  471;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60

Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120

Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
```

```
Db        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
Db        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
Db        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
Db        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
Db        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTGTAG 471
Db        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTGTAG 471
```

RESULT 2
US-10-622-817-10
; Sequence 10, Application US/10622817
; Publication No. US20040214763A1
; GENERAL INFORMATION:
; APPLICANT: CORTI, Olga
; APPLICANT: HAMPE, Cornelia
; APPLICANT: BRICE, Alexis
; APPLICANT: PRADIER, Laurent
; APPLICANT: ROONEY, Thomas
; APPLICANT: FOURNIER, Alain
; TITLE OF INVENTION: METHOD FOR DETERMINING THE ABILITY OF A COMPOUND TO MODIFY THE
; TITLE OF INVENTION: INTERACTION BETWEEN PARKIN AND THE P38 PROTEIN
; FILE REFERENCE: FRAV2002-0020 US NP
; CURRENT APPLICATION NUMBER: US/10/622,817
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: GB 0229934.5
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/396,929
; PRIOR FILING DATE: 2002-07-18
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(471)
US-10-622-817-10

Query Match 100.0%; Score 471; DB 18; Length 471;
Best Local Similarity 100.0%; Pred. No. 9.5e-157;
Matches 471; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
Db          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
Db         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
Db        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
Db        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
Db        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
Db        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
Db        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTGTAG 471
Db        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTGTAG 471
```

RESULT 3
US-10-839-688-12
; Sequence 12, Application US/10839688
; Publication No. US20050014173A1
; GENERAL INFORMATION:
; APPLICANT: Farrer, Matthew J.
; TITLE OF INVENTION: PARKINSON'S DISEASE MARKERS
; FILE REFERENCE: 07039-448001
; CURRENT APPLICATION NUMBER: US/10/839,688
; CURRENT FILING DATE: 2004-05-05
; PRIOR APPLICATION NUMBER: US 60/468,832
; PRIOR FILING DATE: 2003-05-08
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-839-688-12

Query Match 99.0%; Score 466.4; DB 19; Length 2955;
Best Local Similarity 99.8%; Pred. No. 1.1e-154;
Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
Db        499 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 558
Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
Db        559 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 618
Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
Db        619 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 678
Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
Db        679 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 738
Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
Db        739 CACCCCACCTCTGACAAGGAAACACCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 798
Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
Db        799 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 858
Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
Db        859 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 918
Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
```

```
Db        919 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 966


RESULT 4
US-10-473-226-1
; Sequence 1, Application US/10473226
; Publication No. US20040198650A1
; GENERAL INFORMATION:
;  APPLICANT: NsGene A/S
;  TITLE OF INVENTION: Means for inhibiting proteolytical processing of Parkin
;  FILE REFERENCE: 506-204-WO
;  CURRENT APPLICATION NUMBER: US/10/473,226
;  CURRENT FILING DATE:  2003-09-29
;  PRIOR APPLICATION NUMBER: DK PA 2001 00525
;  PRIOR FILING DATE: 2001-03-29
;  PRIOR APPLICATION NUMBER: US 60/281,286
;  PRIOR FILING DATE: 2001-04-03
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 2960
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: mRNA
;   LOCATION: (1)..(2960)
;   OTHER INFORMATION:
US-10-473-226-1

  Query Match             99.0%;  Score 466.4;  DB 18;  Length 2960;
  Best Local Similarity   99.8%;  Pred. No. 1.1e-154;
  Matches  467;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 563

Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 623

Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 683

Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 743

Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        744 CACCCCACCTCTGACAAGGAAACACCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 803

Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 863

Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 923

Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 971


RESULT 5
US-10-473-226-3
; Sequence 3, Application US/10473226
; Publication No. US20040198650A1
; GENERAL INFORMATION:
;  APPLICANT: NsGene A/S
;  TITLE OF INVENTION: Means for inhibiting proteolytical processing of Parkin
;  FILE REFERENCE: 506-204-WO
;  CURRENT APPLICATION NUMBER: US/10/473,226
;  CURRENT FILING DATE:  2003-09-29
;  PRIOR APPLICATION NUMBER: DK PA 2001 00525
;  PRIOR FILING DATE: 2001-03-29
;  PRIOR APPLICATION NUMBER: US 60/281,286
;  PRIOR FILING DATE: 2001-04-03
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 2960
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: mRNA
;   LOCATION: (1)..(2960)
;   OTHER INFORMATION:
US-10-473-226-3

  Query Match             99.0%;  Score 466.4;  DB 18;  Length 2960;
  Best Local Similarity   99.8%;  Pred. No. 1.1e-154;
  Matches  467;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 563

Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 623

Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 683

Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 743

Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        744 CACCCCACCTCTGACAAGGAAACACCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 803

Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 863

Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 923

Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 971


RESULT 6
US-10-622-817-7
; Sequence 7, Application US/10622817
; Publication No. US20040214763A1
; GENERAL INFORMATION:
;  APPLICANT: CORTI, Olga
;  APPLICANT:  HAMPE, Cornelia
;  APPLICANT:  BRICE, Alexis
;  APPLICANT:  PRADIER, Laurent
;  APPLICANT:  ROONEY, Thomas
;  APPLICANT:  FOURNIER, Alain
;  TITLE OF INVENTION: METHOD FOR DETERMINING THE ABILITY OF A COMPOUND TO MODIFY THE
;  TITLE OF INVENTION:  INTERACTION BETWEEN PARKIN AND THE P38 PROTEIN
;  FILE REFERENCE: FRAV2002-0020 US NP
;  CURRENT APPLICATION NUMBER: US/10/622,817
;  CURRENT FILING DATE:  2003-07-18
;  PRIOR APPLICATION NUMBER: GB 0229934.5
```

```
;  PRIOR FILING DATE: 2002-12-20
;  PRIOR APPLICATION NUMBER: US 60/396,929
;  PRIOR FILING DATE: 2002-07-18
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
;   LENGTH: 2960
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (102)..(1499)
US-10-622-817-7

  Query Match           99.0%;  Score 466.4;  DB 18;  Length 2960;
  Best Local Similarity  99.8%;  Pred. No. 1.1e-154;
  Matches  467;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 563

Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 623

Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 683

Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 743

Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        744 CACCCCACCTCTGACAAGGAAACACCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 803

Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 863

Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 923

Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 971


RESULT 7
US-10-776-604-1
; Sequence 1, Application US/10776604
; Publication No. US20050003385A1
; GENERAL INFORMATION:
;  APPLICANT: Shimizu, Nobuyoshi
;  APPLICANT:  Mizuno, Yoshikuni
;  TITLE OF INVENTION: DNAs or Genes Participating in Parkinson's Disease
;  FILE REFERENCE: 0652.2110001
;  CURRENT APPLICATION NUMBER: US/10/776,604
;  CURRENT FILING DATE:  2004-02-12
;  PRIOR APPLICATION NUMBER: 09/601,844
;  PRIOR FILING DATE: 2000-08-09
;  PRIOR APPLICATION NUMBER: PCT/JP99/00545
;  PRIOR FILING DATE: 1999-02-09
;  PRIOR APPLICATION NUMBER: JP 10/27531
;  PRIOR FILING DATE: 1998-02-09
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 2960
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (102)..(1496)
;   OTHER INFORMATION:
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (102)..(108)
;   OTHER INFORMATION: Exon 1
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (109)..(272)
;   OTHER INFORMATION: Exon 2
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (273)..(513)
;   OTHER INFORMATION: Exon 3
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (514)..(635)
;   OTHER INFORMATION: Exon 4
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (636)..(719)
;   OTHER INFORMATION: Exon 5
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (720)..(835)
;   OTHER INFORMATION: Exon 6
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (836)..(972)
;   OTHER INFORMATION: Exon 7
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (973)..(1034)
;   OTHER INFORMATION: Exon 8
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1035)..(1184)
;   OTHER INFORMATION: Exon 9
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1185)..(1268)
;   OTHER INFORMATION: Exon 10
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1269)..(1386)
;   OTHER INFORMATION: Exon 11
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1387)..(2960)
;   OTHER INFORMATION: Exon 12
US-10-776-604-1

  Query Match           99.0%;  Score 466.4;  DB 18;  Length 2960;
  Best Local Similarity  99.8%;  Pred. No. 1.1e-154;
  Matches  467;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 563

Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 623

Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 683

Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        684 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 743

Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        744 CACCCCACCTCTGACAAGGAAACACCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 803

Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 863

Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 923

Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 971


RESULT 8
US-10-839-688-11
; Sequence 11, Application US/10839688
; Publication No. US20050014173A1
; GENERAL INFORMATION:
;  APPLICANT: Farrer, Matthew J.
;  TITLE OF INVENTION: PARKINSON'S DISEASE MARKERS
;  FILE REFERENCE: 07039-448001
;  CURRENT APPLICATION NUMBER: US/10/839,688
;  CURRENT FILING DATE:  2004-05-05
;  PRIOR APPLICATION NUMBER: US 60/468,832
;  PRIOR FILING DATE: 2003-05-08
;  NUMBER OF SEQ ID NOS: 81
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 2960
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-839-688-11

  Query Match            99.0%;  Score 466.4;  DB 19;  Length 2960;
  Best Local Similarity  99.8%;  Pred. No. 1.1e-154;
  Matches  467;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 563

Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 623

Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 683

Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 743

Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        744 CACCCCACCTCTGACAAGGAAACACCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 803

Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 863

Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 923

Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 971


RESULT 9
US-10-622-817-13
; Sequence 13, Application US/10622817
; Publication No. US20040214763A1
; GENERAL INFORMATION:
;  APPLICANT: CORTI, Olga
;  APPLICANT:  HAMPE, Cornelia
;  APPLICANT:  BRICE, Alexis
;  APPLICANT:  PRADIER, Laurent
;  APPLICANT:  ROONEY, Thomas
;  APPLICANT:  FOURNIER, Alain
;  TITLE OF INVENTION: METHOD FOR DETERMINING THE ABILITY OF A COMPOUND TO MODIFY THE
;  TITLE OF INVENTION:  INTERACTION BETWEEN PARKIN AND THE P38 PROTEIN
;  FILE REFERENCE: FRAV2002-0020 US NP
;  CURRENT APPLICATION NUMBER: US/10/622,817
;  CURRENT FILING DATE:  2003-07-18
;  PRIOR APPLICATION NUMBER: GB 0229934.5
;  PRIOR FILING DATE: 2002-12-20
;  PRIOR APPLICATION NUMBER: US 60/396,929
;  PRIOR FILING DATE: 2002-07-18
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
;   LENGTH: 1470
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1470)
US-10-622-817-13

  Query Match            73.7%;  Score 347;  DB 18;  Length 1470;
  Best Local Similarity  83.9%;  Pred. No. 2.7e-112;
  Matches  392;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;

Qy          2 GAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGTC 61
              || ||||||| || || | | | |||| |||||||||| ||| || ||||||||||||||| |
Db        476 GAGGTCCAGAAGCTAAACCCACCTACCACAGCTTTTTTGTCTACTGCAAAGGCCCCTGCC 535

Qy         62 AAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTCA 121
              | |  || ||||| || |||||| | || |||||| |||||||||| || ||||| ||||
Db        536 ACAAGGTCCAGCCTGGGAAACTCCGAGTTCAGTGCGGCACCTGCAGACAAGCAACCCTCA 595

Qy        122 CCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGTG 181
              ||||| ||||||| |||||||||||||||||||| |||||||||||||||||||||||| |
Db        596 CCTTGGCCCAGGGCCCATCTTGCTGGGATGATGTCTTAATTCCAAACCGGATGAGTGGAG 655

Qy        182 AATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCAC 241
              | || ||||| ||| |||||||||||||| || || ||||||||||||||||||||||||
Db        656 AGTGTCAATCTCCAGACTGCCCTGGGACAAGAGCTGAATTTTTCTTTAAATGTGGAGCAC 715

Qy        242 ACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGGA 301
              |||| ||||| |||||||| |||||||||||||||| |||||||| | |  || || || |
Db        716 ACCCAACCTCAGACAAGGACACATCAGTAGCTTTGAACCTGATCACCAACAACAGCCGCA 775

Qy        302 ACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAACT 361
              |||| | |||||  ||||||| || ||||||| ||| ||| |||| ||||| || |||
Db        776 GCATCCCCTGCATCGCGTGCACGGATGTCAGGAACCCTGTCTTGGTCTTCCAATGTAACC 835

Qy        362 CCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGATC 421
              |||||||||||| || || ||||| |||||||| |||||||| ||||| ||||| ||||
Db        836 ACCGCCACGTGATCTGTTTGGACTGCTTCCACTTGTACTGTGTCACAAGGCTCAACGATC 895

Qy        422 GGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              |||||||||| |||||| |||| |||||||||| ||||| ||||||
Db        896 GGCAGTTTGTCCACGACGCTCAGCTTGGCTACTCGCTGCCGTGTGTG 942


RESULT 10
US-10-776-604-3
```

```
; Sequence 3, Application US/10776604
; Publication No. US20050003385A1
; GENERAL INFORMATION:
;  APPLICANT: Shimizu, Nobuyoshi
;  APPLICANT:  Mizuno, Yoshikuni
;  TITLE OF INVENTION: DNAs or Genes Participating in Parkinson's Disease
;  FILE REFERENCE: 0652.2110001
;  CURRENT APPLICATION NUMBER: US/10/776,604
;  CURRENT FILING DATE:  2004-02-12
;  PRIOR APPLICATION NUMBER: 09/601,844
;  PRIOR FILING DATE: 2000-08-09
;  PRIOR APPLICATION NUMBER: PCT/JP99/00545
;  PRIOR FILING DATE: 1999-02-09
;  PRIOR APPLICATION NUMBER: JP 10/27531
;  PRIOR FILING DATE: 1998-02-09
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 2876
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (102)..(1412)
;   OTHER INFORMATION:
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (102)..(108)
;   OTHER INFORMATION: Exon 1
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (109)..(272)
;   OTHER INFORMATION: Exon 2
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (273)..(513)
;   OTHER INFORMATION: Exon 3
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (514)..(635)
;   OTHER INFORMATION: Exon 4
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (636)..(751)
;   OTHER INFORMATION: Exon 6
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (752)..(888)
;   OTHER INFORMATION: Exon 7
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (889)..(950)
;   OTHER INFORMATION: Exon 8
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (951)..(1100)
;   OTHER INFORMATION: Exon 9
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1101)..(1184)
;   OTHER INFORMATION: Exon 10
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1185)..(1302)
;   OTHER INFORMATION: Exon 11
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1303)..(2876)
;   OTHER INFORMATION: Exon 12
US-10-776-604-3

  Query Match           61.2%;  Score 288.4;  DB 18;  Length 2876;
  Best Local Similarity 81.8%;  Pred. No. 2.9e-91;
  Matches  383;  Conservative    0;  Mismatches    1;  Indels   84;  Gaps    1;

Qy          1 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTTATGTGTATTGCAAAGGCCCCTGT 563

Qy         61 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 CAAAGAGTGCAGCCGGGAAAACTCAGGGTACAGTGCAGCACCTGCAGGCAGGCAACGCTC 623

Qy        121 ACCTTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGT 180
              ||||||||||||
Db        624 ACCTTGACCCAG------------------------------------------------ 635

Qy        181 GAATGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCA 240
                                                  ||||||||||||||||||||||||
Db        636 ----------------------------------GAATTTTTCTTTAAATGTGGAGCA 659

Qy        241 CACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 300
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        660 CACCCCACCTCTGACAAGGAAACACCAGTAGCTTTGCACCTGATCGCAACAAATAGTCGG 719

Qy        301 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 AACATCACTTGCATTACGTGCACAGACGTCAGGAGCCCCGTCCTGGTTTTCCAGTGCAAC 779

Qy        361 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 TCCCGCCACGTGATTTGCTTAGACTGTTTCCACTTATACTGTGTGACAAGACTCAATGAT 839

Qy        421 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 468
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        840 CGGCAGTTTGTTCACGACCCTCAACTTGGCTACTCCCTGCCTTGTGTG 887


RESULT 11
US-10-839-688-6
; Sequence 6, Application US/10839688
; Publication No. US20050014173A1
; GENERAL INFORMATION:
;  APPLICANT: Farrer, Matthew J.
;  TITLE OF INVENTION: PARKINSON'S DISEASE MARKERS
;  FILE REFERENCE: 07039-448001
;  CURRENT APPLICATION NUMBER: US/10/839,688
;  CURRENT FILING DATE:  2004-05-05
;  PRIOR APPLICATION NUMBER: US 60/468,832
;  PRIOR FILING DATE: 2003-05-08
;  NUMBER OF SEQ ID NOS: 81
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 650
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-839-688-6

  Query Match           29.4%;  Score 138.6;  DB 19;  Length 650;
  Best Local Similarity 99.3%;  Pred. No. 3.4e-38;
  Matches  138;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        330 CAGGAGCCCCGTCCTGGTTTTCCAGTGCAACTCCCGCCACGTGATTTGCTTAGACTGTTT 389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 CAGGAGCCCCGTCCTGGTTTTCCAGTGCAACTCCCGCCACGTGATTTGCTTAGACTGTTT 267

Qy        390 CCACTTATACTGTGTGACAAGACTCAATGATCGGCAGTTTGTTCACGACCCTCAACTTGG 449
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        268 CCACTTATACTGTGTGACAAGACTCAATGATCGGCAGTTTGTTCACGACCCTCAACTTSG 327

Qy        450 CTACTCCCTGCCTTGTGTG 468
              ||||||||||||||||||
Db        328 CTACTCCCTGCCTTGTGTG 346


RESULT 12
```

US-10-839-688-4
; Sequence 4, Application US/10839688
; Publication No. US20050014173A1
; GENERAL INFORMATION:
; APPLICANT: Farrer, Matthew J.
; TITLE OF INVENTION: PARKINSON'S DISEASE MARKERS
; FILE REFERENCE: 07039-448001
; CURRENT APPLICATION NUMBER: US/10/839,688
; CURRENT FILING DATE: 2004-05-05
; PRIOR APPLICATION NUMBER: US 60/468,832
; PRIOR FILING DATE: 2003-05-08
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 660
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-839-688-4

Query Match 19.1%; Score 89.8; DB 19; Length 660;
Best Local Similarity 91.3%; Pred. No. 8.4e-21;
Matches 94; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 124 TTGACCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGTGAA 183
|| | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 236 TTCCCAAAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCAAACCGGATGAGTGGTGAA 295

Qy 184 TGCCAATCCCCACACTGCCCTGGGACTAGTGCAGAATTTTTCT 226
|||||||||||||||||||||||||||||||||||:| | ||
Db 296 TGCCAATCCCCACACTGCCCTGGGACTAGTGCAGWAAGTACCT 338

RESULT 13
US-09-908-975-26043
; Sequence 26043, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26043
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-908-975-26043

Query Match 10.7%; Score 50.2; DB 10; Length 65;
Best Local Similarity 87.3%; Pred. No. 3.4e-07;
Matches 55; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 227 TTAAATGTGGAGCACACCCCACCTCTGACAAGGAAACATCAGTAGCTTTGCACCTGATCG 286
|||||||||||||||||| ||||| |||||||| || || ||||||||| ||||||||
Db 1 TTAAATGTGGAGCACACCCAACCTCAGACAAGGACACGTCGGTAGCTTTGAACCTGATCA 60

Qy 287 CAA 289
| ||
Db 61 CCA 63

RESULT 14
US-10-776-604-65
; Sequence 65, Application US/10776604
; Publication No. US20050003385A1
; GENERAL INFORMATION:
; APPLICANT: Shimizu, Nobuyoshi
; APPLICANT: Mizuno, Yoshikuni
; TITLE OF INVENTION: DNAs or Genes Participating in Parkinson's Disease
; FILE REFERENCE: 0652.2110001
; CURRENT APPLICATION NUMBER: US/10/776,604
; CURRENT FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 09/601,844
; PRIOR FILING DATE: 2000-08-09
; PRIOR APPLICATION NUMBER: PCT/JP99/00545
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: JP 10/27531
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 65
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-776-604-65

Query Match 7.4%; Score 34.8; DB 18; Length 40;
Best Local Similarity 94.7%; Pred. No. 0.081;
Matches 36; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 128 CCCAGGGTCCATCTTGCTGGGATGATGTTTTAATTCCA 165
|| | ||||||||||||||||||||||||||||||||||
Db 3 CCAAAGGTCCATCTTGCTGGGATGATGTTTTAATTCCA 40

RESULT 15
US-09-735-705-293
; Sequence 293, Application US/09735705
; Patent No. US20020052329A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C14
; CURRENT APPLICATION NUMBER: US/09/735,705
; CURRENT FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 293
; LENGTH: 361
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(361)
; OTHER INFORMATION: n = A,T,C or G
US-09-735-705-293

Query Match 7.4%; Score 34.8; DB 9; Length 361;
Best Local Similarity 59.4%; Pred. No. 0.25;
Matches 57; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 203 CTGGGACTAGTGCAGAATTTTTCTTTAAATGTGGAGCACACCCCACCTCTGACAAGGAAA 262
|| | || || |||||| ||| || || ||| || ||| | | || |||

```
Db        134 CTTGAACACGTCTTCAATTTTCCTTCCAAAATGCTGCATGCCACACTTGAGGTAACGAAG 193

Qy        263 CATCAGTAGCTTTGCACCTGATCGCAACAAATAGTC 298
              ||  ||||  |||  || |||  || |  || | ||
Db        194 CANAAGTATTTTTAAACATGACAGCTAANAACATTC 229
```

Search completed: February 26, 2005, 13:19:16
Job time : 286.738 secs